

```
/ GENERAL INFORMATION:
/ APPLICANT: The Procter & Gamble Company
/ APPLICANT: Peters, Kevin
/ APPLICANT: Thompson, Larry
/ APPLICANT: Wang, Feng
/ APPLICANT: Greis, Kenneth
/ TITLE OF INVENTION: Angiogenesis Modulating Proteins
/ FILE REFERENCE: 8865M
/ CURRENT APPLICATION NUMBER: US/10/316,253
/ CURRENT FILING DATE: 2002-12-10
/ PRIOR APPLICATION NUMBER: US 60/355,295
/ PRIOR FILING DATE: 2002-02-08
/ NUMBER OF SEQ ID NOS: 308
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 118
/ LENGTH: 98
/ TYPE: PRT
/ ORGANISM: Sus scrofa
/ US-10-316-253-118

Query Match      73.9%; Score 17; DB 29; Length 98;
Best Local Similarity 25.0%; Pred. No. 6.2e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 WXXXXXXG 8
      |
Db      81 WATTSTSG 88
```

Search completed: August 16, 2003, 14:42:27
Job time : 378 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:35:04 ; Search time 22 Seconds
(without alignments)
9.375 Million cell updates/sec

Title: US-09-185-908-1

Perfect score: 23

Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 121698 seqs, 25781089 residues

Total number of hits satisfying chosen parameters: 121698

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA_New:*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	73.9	109	6	US-10-273-573-7363
2	17	73.9	964	6	US-10-292-798-622
3	16	69.6	103	6	US-10-273-573-5865
4	16	69.6	202	7	US-60-490-890-1601
5	16	69.6	229	6	US-10-286-897-2671
6	16	69.6	229	6	US-10-258-898A-2671
7	16	69.6	232	6	US-10-603-113-19116
8	16	69.6	268	7	US-60-487-610-2724
9	16	69.6	268	7	US-60-487-610-2725
10	16	69.6	283	6	US-10-627-476-588
11	16	69.6	298	6	US-10-286-897-6243
12	16	69.6	298	6	US-10-258-898A-6243
13	16	69.6	361	6	US-10-460-614-50
14	16	69.6	386	6	US-10-603-113-15327
15	16	69.6	395	6	US-10-293-244-3071
16	16	69.6	397	6	US-10-627-476-628
17	16	69.6	401	7	US-60-478-196-3257
18	16	69.6	414	6	US-10-603-114-7982
19	16	69.6	418	6	US-10-617-320-3619
20	16	69.6	433	7	US-60-479-073-106
21	16	69.6	489	6	US-10-286-897-6797
22	16	69.6	469	6	US-10-258-898A-6797
23	16	69.6	510	6	US-10-613-520-1065
24	16	69.6	515	7	US-60-478-196-3014
25	16	69.6	517	6	US-10-286-897-5929
26	16	69.6	517	6	US-10-258-898A-5929

27	16	69.6	590	6	US-10-612-783-4715	Sequence 4715, Ap
28	16	69.6	598	6	US-10-612-783-4252	Sequence 4252, Ap
29	16	69.6	598	6	US-10-612-783-4253	Sequence 4253, App
30	16	69.6	629	6	US-10-627-476-624	Sequence 624, App
31	16	69.6	686	6	US-10-603-114-8253	Sequence 8253, Ap
32	16	69.6	724	6	US-10-612-783-6949	Sequence 6949, Ap
33	16	69.6	743	6	US-10-374-780A-2670	Sequence 2670, Ap
34	16	69.6	772	6	US-10-613-520-1538	Sequence 1538, Ap
35	16	69.6	772	6	US-10-612-783-5950	Sequence 6950, Ap
36	16	69.6	774	7	US-60-487-610-1561	Sequence 1561, Ap
37	16	69.6	788	7	US-60-478-196-3303	Sequence 3303, Ap
38	16	69.6	819	6	US-10-613-520-1486	Sequence 1486, Ap
39	16	69.6	819	6	US-10-613-783-7015	Sequence 7015, Ap
40	16	69.6	966	6	US-10-286-897-1824	Sequence 1824, Ap
41	16	69.6	966	6	US-10-258-898A-1824	Sequence 1824, Ap
42	16	69.6	977	6	US-10-603-113-15579	Sequence 15579, A
43	16	69.6	1013	6	US-10-286-897-1823	Sequence 1823, Ap
44	16	69.6	1013	6	US-10-258-898A-1823	Sequence 1823, Ap
45	16	69.6	1025	6	US-10-286-897-1825	Sequence 1825, Ap

ALIGNMENTS

RESULT 1

US-10-273-573-7363
; Sequence 7363, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 05/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7363
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(109)
; OTHER INFORMATION: Xaa = X or * as defined in table 2
US-10-273-573-7363

Query Match 73.9%; Score 17; DB 6; Length 109;
Best Local Similarity 25.0%; Pred. No. 9.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 19 WASSASSG 26

RESULT 2

US-10-292-798-622
; Sequence 622, Application US/10292798
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789

; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 622
; LENGTH: 964
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-622

Query Match 73.9%; Score 17; DB 6; Length 964;
Best Local Similarity 25.0%; Pred. No. 4.9e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 241 WTSASTSG 248

RESULT 3

US-10-273-573-5865
; Sequence 5865, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 5865
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (35)...(52)
; OTHER INFORMATION: BETA HAEMOGLOBIN SIGNATURE domain identified by eMAPPIX,
; OTHER INFORMATION: accession number PR00814A, p-value=9.000e-09, raw score of 12.94
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (9)...(45)
; OTHER INFORMATION: Globin domain identified by Pfam, accession name globin, E-
; OTHER INFORMATION: value=0.0024, Pfam score of 16.1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(103)
; OTHER INFORMATION: Xaa - X or * as defined in Table 2
US-10-273-573-5865

Query Match 69.6%; Score 16; DB 6; Length 103;
Best Local Similarity 25.0%; Pred. No. 1.5e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 85 WADASAG 92

RESULT 4

US-60-490-890-1601
; Sequence 1601, Application US/60490890
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 PSP

; CURRENT APPLICATION NUMBER: US/60/490,890
; CURRENT FILING DATE: 2003-07-29
; NUMBER OF SEQ ID NOS: 2779
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1601
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-490-890-1601

Query Match 69.6%; Score 16; DB 7; Length 202;
Best Local Similarity 25.0%; Pred. No. 2.5e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 25 WASITTTG 32

RESULT 5

US-10-286-897-2671
; Sequence 2671, Application US/10286897,
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 2671
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-897-2671

Query Match 69.6%; Score 16; DB 6; Length 229;
Best Local Similarity 25.0%; Pred. No. 2.8e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 52 WASITTTG 59

RESULT 6

US-10-258-898A-2671
; Sequence 2671, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/258,898A
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317

```

; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 2671
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-898A-2671

```

```

Query Match          69.6%; Score 16; DB 6; Length 229;
Best Local Similarity 25.0%; Pred. No. 2.8e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY      1 WXXXXXXG 8
      |
DB      52 WASTITTG 59

```

RESULT 7

```

US-10-603-113-19116
; Sequence 19116, Application US/10603113
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/10/603,113
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/248,796
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 19116
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-603-113-19116

```

```

Query Match          69.6%; Score 16; DB 6; Length 232;
Best Local Similarity 25.0%; Pred. No. 2.8e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY      1 WXXXXXXG 8
      |
DB      33 WKASASSG 40

```

RESULT 8

```

US-60-487-610-2724
; Sequence 2724, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2724

```

```

; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-2724

```

```

Query Match          69.6%; Score 16; DB 7; Length 268;
Best Local Similarity 25.0%; Pred. No. 3.1e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY      1 WXXXXXXG 8
      |
DB      84 WTDATSG 91

```

RESULT 9

```

US-60-487-610-2725
; Sequence 2725, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2725
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-487-610-2725

```

```

Query Match          69.6%; Score 16; DB 7; Length 268;
Best Local Similarity 25.0%; Pred. No. 3.1e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY      1 WXXXXXXG 8
      |
DB      84 WTDATSG 91

```

RESULT 10

```

US-10-627-476-588
; Sequence 588, Application US/10627476
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schoeder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-125PCPN
; CURRENT APPLICATION NUMBER: US/10/627,476
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 09/602,787
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6

```



```

; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 588
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-627-476-588

Query Match          69.6%; Score 16; DB 6; Length 283;
Best Local Similarity 25.0%; Pred. No. 3.3e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
Db 248 WHASTAAG 255

RESULT 11
US-10-286-897-6243
; Sequence 6243, Application US/10286897
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 6243
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-897-6243

Query Match          69.6%; Score 16; DB 6; Length 298;
Best Local Similarity 25.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
Db 121 WASTITG 128

RESULT 12
US-10-258-898A-6243
; Sequence 6243, Application US/10258898A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/258,898A
; CURRENT FILING DATE: 2002-10-29

```

```

; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 6243
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-898A-6243

Query Match          69.6%; Score 16; DB 6; Length 298;
Best Local Similarity 25.0%; Pred. No. 3.4e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
Db 121 WASTITG 128

RESULT 13
US-10-460-614-50
; Sequence 50, Application US/10460614
; GENERAL INFORMATION:
; APPLICANT: Nabil Elshourbagy
; APPLICANT: Derk Jon Bergsma
; APPLICANT: Catherine E. Ellis
; APPLICANT: Wendy S. Fuetterer
; APPLICANT: Mahanandeeswar Gattu
; APPLICANT: Israel Gloger
; APPLICANT: Stephanie F. Guerrero
; APPLICANT: Jeffrey Hill
; APPLICANT: Pamela A. Lane
; APPLICANT: Roberto Anibal Macina
; APPLICANT: Andrew Medhurst
; APPLICANT: David Michalovich
; APPLICANT: Jeffrey L. Mooney
; APPLICANT: Menelas N. Pangalos
; APPLICANT: Ganesh Madhusudan Sathe
; APPLICANT: Usman Shabon
; APPLICANT: Melanie Stammers
; APPLICANT: Ping Tsui
; APPLICANT: Lisa Vawter
; APPLICANT: Shuji Hinuma
; APPLICANT: Shoji Fukusumi
; APPLICANT: Yuji Kawamata
; APPLICANT: Erin M. Toland
; APPLICANT: Yuan Zhu
; APPLICANT: Xiaotong Li
; TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
; FILE REFERENCE: GP-70744B-C6
; CURRENT APPLICATION NUMBER: US/10/460,614
; CURRENT FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: 10/309,870
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 10/185,465
; PRIOR FILING DATE: 2002-07-27
; PRIOR APPLICATION NUMBER: 09/442,367
; PRIOR FILING DATE: 2001-08-29

```

; PRIOR APPLICATION NUMBER: 09/841,736
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 361
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-460-614-50

Query Match 69.6%; Score 16; DB 6; Length 361;
Best Local Similarity 25.0%; Pred. No. 3.9e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 95 WAAAAALG 102

RESULT 14

US-10-603-113-15327
; Sequence 15327, Application US/10603113
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/10/603,113
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/248,796
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 15327
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-603-113-15327

Query Match 69.6%; Score 16; DB 6; Length 386;
Best Local Similarity 25.0%; Pred. No. 4.1e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 194 WSGATATG 201

RESULT 15

US-10-293-244-3071
; Sequence 3071, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; PRIOR FILING DATE: 2002-11-12
; CURRENT FILING DATE: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 3071
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-3071

Query Match 69.6%; Score 16; DB 6; Length 395;
Best Local Similarity 25.0%; Pred. No. 4.2e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 31 WLTSSAAG 38

Search completed: August 16, 2003, 14:43:58
Job time : 23 secs

GenCore version 5.1.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:31:20 ; Search time 29 Seconds
(without alignments)
11.672 Million cell updates/sec

Title: US-09-185-908-1

Perfect score: 23

Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2.6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	73.9	137	4	US-09-252-991A-32654
2	17	73.9	146	4	US-09-252-991A-20461
3	17	73.9	148	4	US-09-252-991A-17476
4	17	73.9	155	4	US-09-252-991A-17807
5	17	73.9	229	4	US-09-252-991A-29247
6	17	73.9	239	4	US-09-252-991A-19211
7	17	73.9	297	4	US-09-252-991A-28842
8	17	73.9	303	4	US-09-252-991A-28605
9	17	73.9	527	3	US-08-907-229-2
10	17	73.9	527	4	US-09-574-462-2
11	17	73.9	527	4	US-08-677-970-3
12	17	73.9	593	2	US-08-468-249A-21
13	16	69.6	17	4	US-09-595-344-20
14	16	69.6	20	4	US-09-595-344-18
15	16	69.6	41	4	US-09-194-139-4
16	16	69.6	78	6	5514590-7
17	16	69.6	100	4	US-09-252-991A-24179
18	16	69.6	105	3	US-09-001-472-2
19	16	69.6	105	4	US-08-857-389-26
20	16	69.6	105	4	US-09-775-827A-2
21	16	69.6	105	4	US-09-775-827A-13
22	16	69.6	111	4	US-09-252-991A-23578
23	16	69.6	120	4	US-09-252-991A-21544
24	16	69.6	127	4	US-09-252-991A-26303
25	16	69.6	129	4	US-09-732-210-63
26	16	69.6	130	4	US-09-732-210-1130
27	16	69.6	130	4	US-09-732-210-1135

Sequence 1139, Ap
Sequence 1315, Ap
Sequence 1316, Ap
Sequence 18434, A
Sequence 1141, Ap
Sequence 6111, Ap
Sequence 30960, A
Sequence 25419, A
Sequence 28685, A
Sequence 16705, A
Sequence 26930, A
Sequence 7, Appli
Sequence 28089, A
Sequence 21913, A
Sequence 27890, A
Sequence 27017, A
Sequence 20832, A
Sequence 24725, A

ALIGNMENTS

RESULT 1

US-09-252-991A-32654
; Sequence 32654, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONA
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32654
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32654

Query Match 73.9%; Score 17; DB 4; Length 137;

Best local similarity 25.0%; Pred. No. 3.4e+03; Mismatches 6; Indels 0; Gaps 0;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
| | |
Db 59 WTSTSSG 66

RESULT 2

US-09-252-991A-20461
; Sequence 20461, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMON
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20461
; LENGTH: 146
; TYPE: PRT

```

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20461

Query Match      73.9%; Score 17; DB 4; Length 146;
Best Local Similarity 25.0%; Pred. No. 3.6e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
   | | | | |
Db 69 WSSSTTG 76

RESULT 3
US-09-252-991A-17476
; Sequence 17476, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17476
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17476

Query Match      73.9%; Score 17; DB 4; Length 148;
Best Local Similarity 25.0%; Pred. No. 3.6e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
   | | | | |
Db 52 WTAASSG 59

RESULT 4
US-09-252-991A-17807
; Sequence 17807, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17807
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17807

Query Match      73.9%; Score 17; DB 4; Length 155;
Best Local Similarity 25.0%; Pred. No. 3.7e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
   | | | | |
Db 122 WSAATTG 129

```

```

RESULT 5
US-09-252-991A-29247
; Sequence 29247, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29247
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29247

Query Match      73.9%; Score 17; DB 4; Length 229;
Best Local Similarity 25.0%; Pred. No. 4.9e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
   | | | | |
Db 44 WTSAAASG 51

RESULT 6
US-09-252-991A-19211
; Sequence 19211, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19211
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19211

Query Match      73.9%; Score 17; DB 4; Length 239;
Best Local Similarity 25.0%; Pred. No. 5.1e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
   | | | | |
Db 155 WSSASASG 162

RESULT 7
US-09-252-991A-28842
; Sequence 28842, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```

Query Match 73.9% Score 17; DB 4; Length 297;
Best Local Similarity 25.0%; Pred. No. 5.9e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 28842
LENGTH: 297
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28842

Query Match 73.9% Score 17; DB 4; Length 297;
Best Local Similarity 25.0%; Pred. No. 5.9e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 56 WASSATSG 63

RESULT 8

US-09-252-991A-28605
; Sequence 28605, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28605
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28605

Query Match 73.9% Score 17; DB 4; Length 303;
Best Local Similarity 25.0%; Pred. No. 6e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 139 WSAASTG 146

RESULT 9

US-08-907-229-2
; Sequence 2, Application US/08907229A
; Patent No. 6072048
; GENERAL INFORMATION:
; APPLICANT: Riley, Lee W.
; TITLE OF INVENTION: DNA MOLECULE ENCODING FOR CELLULAR UPTAKE OF
; FILE REFERENCE: 19603/1531
; CURRENT APPLICATION NUMBER: US/08/907,229A
; CURRENT FILING DATE: 1997-08-06
; EARLIER APPLICATION NUMBER: 60/040,097
; EARLIER FILING DATE: 1997-03-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-08-907-229-2

Query Match 73.9% Score 17; DB 3; Length 527;
Best Local Similarity 25.0%; Pred. No. 8.8e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 17 WSAATATG 24

RESULT 10

US-09-574-462-2
; Sequence 2, Application US/09574462
; Patent No. 6399764
; GENERAL INFORMATION:
; APPLICANT: Riley, Lee W.
; TITLE OF INVENTION: DNA MOLECULE ENCODING FOR CELLULAR UPTAKE OF
; FILE REFERENCE: 19603/1531
; CURRENT APPLICATION NUMBER: US/09/574,462
; CURRENT FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 08/907,229
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: 60/040,097
; PRIOR FILING DATE: 1997-03-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-574-462-2

Query Match 73.9% Score 17; DB 4; Length 527;
Best Local Similarity 25.0%; Pred. No. 8.8e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 17 WSAATATG 24

RESULT 11

US-08-677-970-3
; Sequence 3, Application US/08677970
; Patent No. 6444444
; GENERAL INFORMATION:
; APPLICANT: Anand, Naveen N.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: GENES ENCODING MYCOBACTERIAL PROTEINS
; TITLE OF INVENTION: ASSOCIATED WITH CELL BINDING AND CELL ENTRY AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,970
; FILING DATE: 10-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-592 MIS:jbb
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-677-970-3

Query Match 73.9%; Score 17; DB 2; Length 527;
Best Local Similarity 25.0%; Pred. No. 8.6e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 17 WSAATATG 24

RESULT 12
US-08-468-249A-21
Sequence 21, Application US/08468249A
Patent No. 5886148
GENERAL INFORMATION:

APPLICANT: Segre et al., Gino V.
TITLE OF INVENTION: PARATHYROID HORMONE RECEPTOR AND DNA
TITLE OF INVENTION: ENCODING SAME
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08468,249A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,475
FILING DATE: 06-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/681,702
FILING DATE: 04-MAY-1991

ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00786/071003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906

INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 593 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-249A-21

Query Match 73.9%; Score 17; DB 2; Length 593;
Best Local Similarity 25.0%; Pred. No. 9.6e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 69 WTSASTSG 76

RESULT 13
US-09-595-344-20
Sequence 20, Application US/09595344
Patent No. 6534286
GENERAL INFORMATION:

APPLICANT: Li, Xin-Liang
TITLE OF INVENTION: Protein Production in Aureobasidium pullulans
FILE REFERENCE: 34-00
CURRENT APPLICATION NUMBER: US/09/595,344
CURRENT FILING DATE: 2000-06-15
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 17
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:amino acid
OTHER INFORMATION: sequence encoded by multiple cloning site linker
OTHER INFORMATION: sequence of SEQ ID NO:21.
US-09-595-344-20

Query Match 69.6%; Score 16; DB 4; Length 17;
Best Local Similarity 25.0%; Pred. No. 1.3e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 4 WSAALG 11

RESULT 14
US-09-595-344-18
Sequence 18, Application US/09595344
Patent No. 6534286
GENERAL INFORMATION:

APPLICANT: Li, Xin-Liang
TITLE OF INVENTION: Protein Production in Aureobasidium pullulans
FILE REFERENCE: 34-00
CURRENT APPLICATION NUMBER: US/09/595,344
CURRENT FILING DATE: 2000-06-15
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 18
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:amino acid
OTHER INFORMATION: sequence encoded by multiple cloning site of SEQ
OTHER INFORMATION: ID NO:19.
US-09-595-344-18

Query Match 69.6%; Score 16; DB 4; Length 20;
Best Local Similarity 25.0%; Pred. No. 1.5e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 7 WSAALG 14

RESULT 15
US-09-194-139-4
Sequence 4, Application US/09194139A
Patent No. 6306608
GENERAL INFORMATION:

APPLICANT: Arai, Takao
TITLE OF INVENTION: ANTI-HUMAN LECT2 ANTIBODY, CELLS PRODUCING THE SAME,
TITLE OF INVENTION: AND METHOD AND KIT FOR ASSAYING THE SAME

; FILE REFERENCE: US application 09/194,139
; CURRENT APPLICATION NUMBER: US/09/194,139A
; CURRENT FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: JP 8-132160
; EARLIER FILING DATE: 1996-05-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer
; OTHER INFORMATION: sequence
US-09-194-139-4

Query Match 69.6%; Score 16; DB 4; Length 41;
Best Local Similarity 25.0%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 WXXXXXXG 8
|
Db 27 WGTSTATG 34

Search completed: August 16, 2003, 14:36:03
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:43:34 ; Search time 38 Seconds
(without alignments)
20.246 Million cell updates/sec

Title: US-09-185-908-1

Perfect score: 23

Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 604

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	47.8	3	3 F37196	bradykinin-potenti
2	11	47.8	4	2 A34626	RPCB-related neuro
3	11	47.8	4	2 B53284	T-cell receptor be
4	11	47.8	4	2 PT0661	T-cell receptor be
5	11	47.8	5	2 A32516	cholecystokinin-5
6	11	47.8	5	2 A60803	neuropeptide - sea
7	11	47.8	5	2 JH0253	gut pentapeptide -
8	11	47.8	5	2 G37196	bradykinin-potenti
9	11	47.8	5	2 PT0281	Ig heavy chain CRD
10	11	47.8	5	2 PT0308	Ig heavy chain CRD
11	11	47.8	5	2 PT0729	T-cell receptor be
12	11	47.8	5	2 PT0380	T-cell receptor be
13	11	47.8	6	2 S66195	alcohol dehydrogen
14	11	47.8	6	2 B34835	dnaA protein - Pse
15	11	47.8	6	2 A31263	dihydrofolate redu
16	11	47.8	6	2 B31263	dihydrofolate redu
17	11	47.8	6	2 A61068	locustakinin - mig
18	11	47.8	6	2 B35640	cerebellar degener
19	11	47.8	6	2 PT0629	T-cell receptor be
20	11	47.8	6	2 PT0532	T-cell receptor be
21	11	47.8	6	2 PT0519	T-cell receptor be
22	11	47.8	6	2 PT0637	T-cell receptor be
23	11	47.8	6	2 PT0641	T-cell receptor be
24	11	47.8	6	2 PT0726	T-cell receptor be
25	11	47.8	6	2 F41946	T-cell receptor ga
26	11	47.8	6	2 PD0028	pev-kinin 2 - pena
27	11	47.8	6	4 I79564	hypothetical TCL3
28	11	47.8	7	2 S21230	dermorphin (trp-4,
29	11	47.8	7	2 A58512	venom heptapeptide

30 11 47.8 7 2 A61081 tryptophyllin, bas
31 11 47.8 7 2 S57274 triacylglycerol li
32 11 47.8 7 2 S09652 hypothetical prote
33 11 47.8 7 2 P00727 H2 class I protein
34 11 47.8 7 2 S33244 neuromodulatory pe
35 11 47.8 7 2 S33245 neuromodulatory pe
36 11 47.8 7 2 S33246 neuromodulatory pe
37 11 47.8 7 2 E48394 glycoprotein compo
38 11 47.8 7 2 PH1602 Ig H chain V-D-J r
39 11 47.8 7 2 PT0526 T-cell receptor be
40 11 47.8 7 2 PT0628 T-cell receptor be
41 11 47.8 7 2 PT0642 T-cell receptor be
42 11 47.8 7 2 PT0722 T-cell receptor be
43 11 47.8 7 2 PT0688 T-cell receptor be
44 11 47.8 7 2 PT0586 T-cell receptor be
45 11 47.8 7 2 PT0728 T-cell receptor be

ALIGNMENTS

RESULT 1

F37196

bradykinin-potentiating peptide 6 - island jararaca

C:Species: Bothrops insularis (island jararaca)

C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C:Accession: F37196

J:Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.

J:Protein Chem. 9, 221-227, 1990

A:Title: Primary structure and biological activity of bradykinin potentiating peptid

A:Reference number: A37196; MUID:90351557; PMID:2386615

A:Accession: F37196

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-3 <CIN>

A:Keywords: proglutamic acid

F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 47.8%; Score 11; DB 3; Length 3;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1

Db 3 W 3

RESULT 2

A34626

RPCB-related neuropeptide - ferruginous spindle

C:Species: Fusinus ferrugineus (ferruginous spindle)

C>Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 31-Dec-1993

C:Accession: A34626

R:Kuroki, Y.; Kanda, T.; Kubota, I.; Fujisawa, Y.; Ikeda, T.; Miura, A.; Minamitake,

Biochem. Biophys. Res. Commun. 167, 273-279, 1990

A:Title: A molluscan neuropeptide related to the crustacean hormone, RPCH.

A:Reference number: A34626; MUID:90179762; PMID:2310394

A:Accession: A34626

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-4 <KUR>

C:Keywords: neuropeptide

Query Match 47.8%; Score 11; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1

Db 4 W 4

RESULT 3

B53284
T-cell receptor beta 2 chain D region, Dbeta2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: B53284
R:Harindranath, N.; Alexander, C.B.; Mage, R.G.
Mol. Immunol. 28, 881-888, 1991
A:Title: Evolutionarily conserved organization and sequences of germline diversity and
A:Reference number: A53284; MUID:91342695; PMID:1678859
A:Accession: B53284
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-4 <HAR>
A:Cross-references: GB:S60737; NID:g233916; PID:g233918
A>Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBI:P:60738)
C:Keywords: T-cell receptor

Query Match 47.8%; Score 11; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
DB 2 W 2

RESULT 4
PT0661
T-cell receptor beta chain V-D-J region (121-LBV) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0661
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0661
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-4 <FEE>
A:Experimental source: day 4 postnatal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 47.8%; Score 11; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
DB 3 W 3

RESULT 5
A32516
cholecystokinin-5 - dog
N:Alternate names: CCK-5
C:Species: Canis lupus familiaris (dog)
C:Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 18-Aug-2000
C:Accession: A32516
R:Shively, J.; Reeve Jr., J.R.; Eysselein, V.E.; Ben-Avram, C.; Vigna, S.R.; Walsh, J.H.
Am. J. Physiol. 252, G272-G275, 1987
A:Title: CCK-5: sequence analysis of a small cholecystokinin from canine brain and intest
A:Reference number: A32516; MUID:87153871; PMID:3826354
A:Accession: A32516
A:Molecule type: protein
A:Residues: 1-5 <SHI>
A:Comment: This peptide corresponds to the five carboxyl-terminal residues of cholecyst
C:Superfamily: gastrin
C:Keywords: amidated carboxyl end; neuropeptide
F:5/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 47.8%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;

QY 1 W 1
|
DB 3 W 3

RESULT 6
A60803
neuropeptide - sea anemone (Anthopleura elegantissima)
C:Species: Anthopleura elegantissima
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: A60803
R:Graff, D.; Grimmelikhuijzen, C.J.P.
Brain Res. 442, 354-358, 1988
A:Title: Isolation of <Glu-Ser-Iu-Arg-Trp-NH-2, a novel neuropeptide from sea anemone
A:Reference number: A60803; MUID:88222764; PMID:2897223
A:Accession: A60803
A:Molecule type: protein
A:Residues: 1-5 <GRA>
C:Keywords: amidated carboxyl end; neuropeptide; pyrrolutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:5/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 47.8%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
DB 5 W 5

RESULT 7
JH0253
gut pentapeptide - Japanese eel
C:Species: Anguilla japonica (Japanese eel)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Apr-1995
C:Accession: JH0253
R:Uesaka, T.; Ikeda, T.; Kubota, I.; Muneoka, Y.; Ando, M.
Biochem. Biophys. Res. Commun. 180, 828-832, 1991
A:Title: Structure and function of a pentapeptide isolated from the gut of the eel.
A:Reference number: JH0253; MUID:92062113; PMID:1953755
A:Accession: JH0253
A:Molecule type: protein
A:Residues: 1-5 <UES>
A:Experimental source: gut
C:Comment: This peptide increased basal tone of the circular muscle of the esophagoga
, and of the circular muscle of the gastro-intestinal junction.

Query Match 47.8%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
DB 3 W 3

RESULT 8
G37196
bradykinin-potentiating peptide 7 - island jararaca
C:Species: Bothrops insularis (island jararaca)
C:Date: 14-Feb-1992 #sequence_revision 01-Dec-1992 #text_change 05-Aug-1994
C:Accession: G37196
R:Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.
J. Protein Chem. 9, 221-227, 1990
A:Title: Primary structure and biological activity of bradykinin potentiating peptide
A:Reference number: A37196; MUID:90351557; PMID:2386615
A:Accession: G37196
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <CIN>

C:Keywords: pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 47.8%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
3 W 3
Db

RESULT 9
PT0281
Ig heavy chain CRD3 region (clone 4-91C) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0281
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PT0222; MUID:91108337; PMID:1895102
A:Accession: PT0281
A:Molecule type: DNA
A:Residues: 1-5 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 47.8%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
4 W 4
Db

RESULT 10
PT0308
Ig heavy chain CRD3 region (clone 6-88) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0308
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PT0222; MUID:91108337; PMID:1895102
A:Accession: PT0308
A:Molecule type: DNA
A:Residues: 1-5 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 47.8%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
2 W 2
Db

RESULT 11
PT0729
T-cell receptor beta chain V-D-J region (120-1J) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0640; PT0685; PT0729
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0640

A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FE2>
A:Experimental source: newborn thymus, strain BALB/c, clone 120-1J
A:Accession: PT0685
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <FE2>
A:Experimental source: day 18 fetal thymus, strain BALB/c, clone 154-1C
A:Accession: PT0729
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-5 <FE3>
A:Experimental source: newborn thymus, strain BALB/c, clone 135-1AG
C:Keywords: T-cell receptor

Query Match 47.8%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
5 W 5
Db

RESULT 12
PT0580
T-cell receptor beta chain V-D-J region (159-2B) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0580
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N region
A:Reference number: PT0509; MUID:91277601; PMID:1711558
A:Accession: PT0580
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FE2>
A:Experimental source: day 19 fetal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 47.8%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
4 W 4
Db

RESULT 13
S66195
alcohol dehydrogenase (EC 1.1.1.1) class III low affinity form - cod (Gadus sp.) (fr
C:Species: Gadus sp. (cod)
C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 12-Jun-1998
C:Accession: S66195
R:Hjelmqvist, L.; Hackett, M.; Shafiqat, J.; Danielsson, O.; Iida, J.; Hendrickson, R
FEBS Lett. 367, 237-240, 1995
A:Title: Multiplicity of N-terminal structures of medium-chain alcohol dehydrogenase
enzyme.
A:Reference number: S66191; MUID:95331382; PMID:7607314
A:Accession: S66195
A:Molecule type: protein
A:Residues: 1-6 <HJE>
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 47.8%; Score 11; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1

Db 5 W 5

RESULT 14

B34835
 dnaA protein - Pseudomonas aeruginosa (fragment)
 C:Species: Pseudomonas aeruginosa
 C:Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 08-Oct-1999
 C:Accession: B34835
 R.Yee, T.W.; Smith, D.W.
 Proc. Natl. Acad. Sci. U.S.A. 87, 1278-1282, 1990
 A:Title: Pseudomonas chromosomal replication origins: a bacterial class distinct from Es
 A:Reference number: A34835; PMID:2106132
 A:Accession: B34835
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-6 <IEE>
 A:Cross-references: GB:M30125; NID:g151419; PIDN:AAA25916.1; PID:g151421
 C:Keywords: DNA binding

Query Match 47.8%; Score 11; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
 |
 Db 6 W 6

RESULT 15

A31263
 dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - Plasmodium f
 C:Species: Plasmodium falciparum
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 20-Mar-1996
 C:Accession: A31263
 R.Peterson, D.S.; Walliker, D.; Wellem, T.E.
 Proc. Natl. Acad. Sci. U.S.A. 85, 9114-9118, 1988
 A:Title: Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase
 A:Reference number: A94217; PMID:9057886; PMID:2904149
 A:Accession: A31263
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-6 <PET>
 C:Keywords: methyltransferase; NADP; oxidoreductase

Query Match 47.8%; Score 11; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
 |
 Db 3 W 3

Search completed: August 16, 2003, 14:48:24
 Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:36:09 ; Search time 23 Seconds
(without alignments)
16.357 Million cell updates/sec

Title: US-09-185-908-1
Perfect score: 23
Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 158

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	11	47.8	4	1 OCP3_OCTMI	P58649 octopus min
2	11	47.8	5	1 BPP7_BOTIN	P30425 bothrops in
3	11	47.8	5	1 UF01_MOUSE	P38639 mus musculu
4	11	47.8	6	1 E101_LITRU	P82096 litoria rub
5	11	47.8	6	1 LOK1_LOCTM	P41491 locusta mig
6	11	47.8	7	1 BRP1_CONIM	P58803 conus imper
7	11	47.8	7	1 TPY1_PACDA	P83455 pachymedusa
8	11	47.8	7	1 TY51_LITRU	P82065 litoria rub
9	11	47.8	7	1 WWA1_ACHFU	P35919 achatina fu
10	11	47.8	7	1 WWA2_ACHFU	P35920 achatina fu
11	11	47.8	7	1 WWA3_ACHFU	P35921 achatina fu
12	11	47.8	8	1 ACL1_THUAL	P18691 thunnus alb
13	11	47.8	8	1 AKHG_GRYBI	P14086 gryllus bim
14	11	47.8	8	1 AKH_LIBAU	P25418 libellula a
15	11	47.8	8	1 AKH_MEML	P25423 melolontha
16	11	47.8	8	1 AKH_TABAT	P14595 tabanus atr
17	11	47.8	8	1 CCKN_MACEU	P30369 macropus eu
18	11	47.8	8	1 COW2_CONPU	P58785 conus purpu
19	11	47.8	8	1 HTF1_PERAM	P04548 periplaneta
20	11	47.8	8	1 HTF2_PERAM	P25419 periplaneta
21	11	47.8	8	1 HTF_TENNO	P25419 tenebrio mo
22	11	47.8	8	1 LCK1_LEUMA	P21140 leucophaea
23	11	47.8	8	1 LCK2_LEUMA	P21141 leucophaea
24	11	47.8	8	1 LCK3_LEUMA	P21142 leucophaea
25	11	47.8	8	1 LCK4_LEUMA	P21143 leucophaea
26	11	47.8	8	1 LCK5_LEUMA	P19987 leucophaea
27	11	47.8	8	1 LCK6_LEUMA	P19988 leucophaea
28	11	47.8	8	1 LCK7_LEUMA	P19989 leucophaea
29	11	47.8	8	1 LCK8_LEUMA	P19990 leucophaea
30	11	47.8	8	1 FLP_BRANA	P81707 brassica na
31	11	47.8	8	1 RPCH_PANBO	P08939 pandalus bo
32	11	47.8	8	1 RT34_BOVIN	P82929 bos taurus
33	6	26.1	3	1 GRWM_HUMAN	P01157 homo sapien

34 6 26.1 4 1 ACH1_ACHFU P35904 achatina fu
35 6 26.1 4 1 DCML_PSECH P19916 pseudomonas
36 6 26.1 4 1 EOS1_HUMAN P02731 homo sapien
37 6 26.1 4 1 OCP1_OCTMI P58649 octopus min
38 6 26.1 5 1 AL14_CARNA P81817 cardinus ma
39 6 26.1 5 1 PAP2_PARNA P81864 pardachirus
40 6 26.1 5 1 RE32_LITRU P82073 litoria rub
41 6 26.1 5 1 TPIS_CANFA P54714 canis famil
42 6 26.1 5 1 UXA4_CHLTR P38005 chlamydia t
43 6 26.1 6 1 CIP1_MYTED P13736 mytilus edu
44 6 26.1 6 1 CIP2_MYTED P13737 mytilus edu
45 6 26.1 6 1 FARP_MONEX P41966 moniezia ex

ALIGNMENTS

RESULT 1
OCP3_OCTMI STANDARD; PRT; 4 AA.
AC P58649;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cardioactive peptides Ocp-3/Ocp-4.
OS Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
OX NCBI_TaxID=89766;
RN [1]
RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=20336815; PubMed=10876044;
RA Wakoshi E., Hisada M., Minakata H.;
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
RT Octopus minor.";
RL Peptides 21:623-630(2000).
CC -!- FUNCTION: Cardioactive; has both positive chronotropic and
CC inotropic effects on the heart. Ocp-4 is a 1000 time less
CC active than ocp-3.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: Ocp-4 has D-Ser instead of L-Ser.
CC -!- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
KW Hormone; D-amino acid.
FT MOD_RES 2 D-SERINE (IN OCP-4).
SQ SEQUENCE 4 AA; 463 MW; 6AB365B810000000 CRC64;
Query Match 47.8%; Score 11; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 W 1

DB 3 W 3

RESULT 2

BPP7_BOTIN STANDARD; PRT; 5 AA.
AC P30425;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bradykinin-potentiating peptide S5.2 (5A) (Angiotensin-converting
DE enzyme inhibitor).
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;

```

RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
RL peptides from Bothrops insularis snake venom.";
RL J. Protein Chem. 9:221-227(1990).
CC -1- FUNCTION: This peptide both inhibits the activity of the
CC angiotensin-converting enzyme and enhances the action of
CC bradykinin by inhibiting the kinases that inactivate it.
CC It acts as an indirect hypotensive agent.
DR PIR: G37196; G37196
KW Hypotensive agent; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 5 AA; 629 MW; 776DC37326B00000 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
Db 3 W 3

RESULT 3
UF01_MOUSE
ID UF01_MOUSE STANDARD; PRT; 5 AA.
AC P38639;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE
RC TISSUE=Fibroblast;
RX MEDLINE=95009907; PubMed=7523108;
RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT "Separation and sequencing of familial and novel murine proteins
RL using preparative two-dimensional gel electrophoresis.";
RL Electrophoresis 15:735-745(1994).
CC -1- MISCELLANEOUS: ON THE 2D-GEL, THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.6, ITS MW IS: 19 KDa.
FT NON_TER 5 5
SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
Db 1 W 1

RESULT 4
E101_LITRU
ID E101_LITRU STANDARD; PRT; 6 AA.
AC P82096;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Electrin 1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE
RC TISSUE=Skin secretion;

```

```

RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:639-645(1999).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD_RES 6 6 AMIDATION.
SQ SEQUENCE 6 AA; 792 MW; 5683704772C9A000 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
Db 5 W 5

RESULT 5
LOK1_LOCM1
ID LOK1_LOCM1 STANDARD; PRT; 6 AA.
AC P41491;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Locustakinin I.
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE
RC TISSUE=Corpora cardiaca;
RX MEDLINE=92262851; PubMed=1585017;
RA Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K.,
RA de Loof A.;
RT "Locustakinin, a novel myotropic peptide from Locusta migratoria,
RT isolation, primary structure and synthesis.";
RL Regul. Pept. 37:49-57(1992).
CC -1- FUNCTION: Myotropic peptide. May be important in the stimulation
CC of ion transport and inhibition of diuretic activity in Malpighian
CC tubules.
CC -1- SUBCELLULAR LOCATION: Secreted.
DR PIR: A61068; A61068.
KW Neuropeptide; Amidation.
FT MOD_RES 6 6 AMIDATION.
SQ SEQUENCE 6 AA; 654 MW; 686365A5B9CDB000 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
Db 5 W 5

RESULT 6
BRHP_CONIM
ID BRHP_CONIM STANDARD; PRT; 7 AA.
AC P38803;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bromheptapeptide Im.
OS Conus imperialis (Imperial cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeconcha; Hypogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=35631;

```

```

RN  SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RP  TISSUE-VEINOM;
RC  MEDLINE=97184108; PubMed=9030520;
RX  Craig A.G., Jimenez E.C., Dykert J., Nielsen D.B., Gulyas J.,
RA  Abogadie F.C., Porter J., Rivier J.E., Cruz L.J., Olivera B.M.,
RA  McIntosh J.M.;
RT  "A novel post-translational modification involving bromination of
RT  tryptophan. Identification of the residue, L-6-bromotryptophan, in
RT  peptides from Conus imperialis and Conus radiatus venom. ";
RL  J. Biol. Chem. 272:4699-4698(1997).
CC  -!- FUNCTION: Does not elicit gross behavioral symptoms when injected
CC  centrally or peripherally in mice.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC  -!- MASS SPECTROMETRY: MW=853.19; METHOD=LSIMS.
DR  PIR; A58512;
KW  Bromination; Amidation; Pyrrolidone carboxylic acid.
FT  DISULFID 2 7
FT  MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT  MOD_RES 6 6 BROMINATION.
FT  MOD_RES 7 7 AMIDATION.
SQ  SEQUENCE 7 AA; 795 MW; 6EA37DC6D87EA6B0 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 6 W 6

RESULT 7
TPFY_PACDA STANDARD; PRT; 7 AA.
AC P83455;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tryptophyllin-1 (Pdt-1).
OS Pachymedusa dactylophora (Giant mexican leaf frog).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
CC Phyllomedusinae; Pachymedusa.
OX NCBI_TaxID=75988;
RN [1]
RP SEQUENCE, MASS SPECTROMETRY, HYDROXYLATION OF PRO-3, AND AMIDATION OF
RP PRO-7.
RC TISSUE-Skin secretion;
RA Chen T.B., Orr D.F., Shaw C.;
RT "Pachymedusa dactylophora tryptophyllin-1 (Pdt-1): structural
RT characterization, pharmacological activity and cloning of precursor
RT cDNA.";
RL Submitted (SEP-2002) to the SWISS-PROT data bank.
CC -!- FUNCTION: Myoactive. Has selective relaxing activity on vascular
CC smooth muscle.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- MASS SPECTROMETRY: MW=809.2; METHOD=MALDI.
DR GO: GO:0005576; C:extracellular; NAS.
DR GO: GO:0045986; P:negative regulation of smooth muscle contractility; NAS.
KW Amphibian defense peptide; Amidation; Hydroxylation.
MOD_RES 3 3 HYDROXYLATION.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 794 MW; 7772D37DC7776350 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 6 W 6

```

```

Db 5 W 5

RESULT 8
TY51_LITRU STANDARD; PRT; 7 AA.
ID TY51_LITRU
AC P82065;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tryptophyllin 5.1.
OS Litoria rubella (Desert tree frog).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
CC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-Skin secretion;
RA Steinbörner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RT 'Litoria rubella'. The skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -!- FUNCTION: May act as a neuromodulator or neurotransmitter.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC -!- MASS SPECTROMETRY: MW=965; METHOD=FAE.
KW Amphibian defense peptide; Amidation; Neuropeptide;
KW Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 983 MW; 7401E9D3676046B0 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 4 W 4

RESULT 9
WWAL_ACHFU STANDARD; PRT; 7 AA.
ID WWAL_ACHFU
AC P35919;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Wwamide-1.
OS Achatina fulica (Giant African snail).
CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
CC Sigamurethra; Achatinidae; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE-Ganglion;
RX MEDLINE=93265912; PubMed=8495720;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Momoto K.;
RT "Wwamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT ganglia of the African giant snail, Achatina fulica.";
RL FEBS Lett. 323:104-108(1993).
CC -!- FUNCTION: EXHIBITS MODULATORY EFFECTS ON THE PERIPHERAL NERVOUS
CC SYSTEM. INHIBITS ACTIVITY ON A CENTRAL NEURON.
CC PIR: S33245; S33245.
DR Neuropeptide; Amidation.
KW Neuropeptide; Amidation.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 993 MW; 7362D5B69B041310 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;

```

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 1 W 1

RESULT 10

WAA2_ACHFU
ID WAA2_ACHFU STANDARD; PRT; 7 AA.
AC P35920; (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Wwamide-2.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion;
RX MEDLINE=93265912; PubMed=8495720;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT ganglia of the African giant snail, Achatina fulica."
RL FEBS Lett. 323:104-108(1993).
DR PIR: S33246; S33246.
KW Neuropeptide; Amidation.
FT MOD_RES 7
SQ SEQUENCE 7 AA; 964 MW; 7362D5B686D32310 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 1 W 1

RESULT 11

WAA3_ACHFU
ID WAA3_ACHFU STANDARD; PRT; 7 AA.
AC P35921;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Wwamide-3.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion;
RX MEDLINE=93265912; PubMed=8495720;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT "Wwamide-1, -2 and -3: novel neuromodulatory peptides isolated from
ganglia of the African giant snail, Achatina fulica."
RL FEBS Lett. 323:104-108(1993).
DR PIR: S33244; S33244.
KW Neuropeptide; Amidation.
FT MOD_RES 7
SQ SEQUENCE 7 AA; 965 MW; 7362D5B69B132310 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 1 W 1

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 12

ACL_THUAL
ID ACL_THUAL STANDARD; PRT; 8 AA.
AC P18691;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Angiotensin-converting enzyme inhibitor.
OS Thunus albacares (Yellowfin tuna) (Neothunnus macropterus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8236;
RN [1]
RP SEQUENCE.
RC TISSUE=Muscle;
RX MEDLINE=88326322; PubMed=3415688;
RA Kohama Y., Matsumoto S., Oka H., Teramoto T., Okabe M., Mimura T.;
RT "Isolation of angiotensin-converting enzyme inhibitor from tuna
muscle."
RL Biochem. Biophys. Res. Commun. 155:332-337(1988).
DR PIR: A31570; A31570.
SQ SEQUENCE 8 AA; 953 MW; 6AA863733051F1B7 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 6 W 6

RESULT 13

AKHG_GRYBI
ID AKHG_GRYBI STANDARD; PRT; 8 AA.
AC P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone G (AKH-G) (RO II).
OS Gryllus bimaculatus (Two-spotted cricket), and
Romalea microptera (lubber grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Ensifera; Gryllidae; Gryllinae;
OC Gryllus.
OX NCBI_TaxID=6999, 7007;
RN [1]
RP SEQUENCE.
RC SPECIES=G.bimaculatus; TISSUE=Corpora cardiaca;
RX MEDLINE=88106553; PubMed=3426616;
RA Gaede G., Rinehart K.L. Jr.;
RT "Primary sequence analysis by fast atom bombardment mass spectrometry
of a peptide with adipokinetic activity from the corpora cardiaca of
the cricket Gryllus bimaculatus."
RL Biochem. Biophys. Res. Commun. 149:908-914(1987).
RN [2]
RP SEQUENCE.
RC SPECIES=R.microptera; TISSUE=Corpora cardiaca;
RX MEDLINE=89145002; PubMed=3226948;
RA Gaede G., Hilbich C., Beyreuther K., Rinehart K.L. Jr.;
RT "Sequence analyses of two neuropeptides of the AKH/RPCH-family from
the lubber grasshopper, Romalea microptera."
RL Peptides 9:681-688(1988).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / RPCH / RPCH FAMILY.
DR PIR: A28004; A28004.

```

DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
Db 8 W 8

RESULT 14
AKH_LIBAU STANDARD; PRT; 8 AA.
AC P25418;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone (AKH).
OS Libellula auripennis (Skimmer dragonfly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Palaeoptera; Odonata; Anisoptera; Libellulidae; Libellula.
OX NCBI_TaxID=9366;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=90359055; PubMed=2390213;
RA Gaede G.;
RT "The putative ancestral peptide of the adipokinetic/red-pigment-
RT concentrating hormone family isolated and sequenced from a
RT dragonfly.";
RL Biol. Chem. Hoppe-Seyler 371:475-483(1990).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; S10596;
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 978 MW; 8665A771A9C452D6 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
Db 8 W 8

RESULT 15
AKH_MEML STANDARD; PRT; 8 AA.
AC P25423;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone (AKH).
OS Melolontha melolontha (Cockchafer),
OS Geotrupes stercorosus (Dor beetle), and
OS Pachnoda marginata (Flower beetle).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Scarabaeiformia;
OC Scarabaeidae; Melolonthinae; Melolontha.

```

```

OX NCBI_TaxID=7061, 7087, 7058;
RN [1]
RP SEQUENCE.
RC SPECIES=M.melolontha, and G.stercorosus; TISSUE=Corpora cardiaca;
RX MEDLINE=91248100; PubMed=2039445;
RA Gaede G.;
RT "A unique charged tyrosine-containing member of the adipokinetic
RT hormone/red-pigment-concentrating hormone peptide family isolated and
RT sequenced from two beetle species.";
RL Biochem. J. 275:671-677(1991).
RN [2]
RP SEQUENCE.
RC SPECIES=P.margaritana; TISSUE=Corpora cardiaca;
RX MEDLINE=92265187; PubMed=1586453;
RA Gaede G., Lopata A., Kellner R., Rinehart K.L. Jr.;
RT "Primary structures of neuropeptides isolated from the corpora
RT cardiaca of various cetonid beetle species determined by
RT pulsed-liquid phase sequencing and tandem fast atom bombardment mass
RT spectrometry.";
RL Biol. Chem. Hoppe-Seyler 373:133-142(1992).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A58641; A58641.
DR PIR; S15422; S15422.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 1022 MW; 867AB775AB544736 CRC64;

Query Match 47.8%; Score 11; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
Db 8 W 8

Search completed: August 16, 2003, 14:45:57
Job time : 25 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:42:34 ; Search time 95 Seconds
(without alignments)
21.731 Million cell updates/sec

Title: US-09-185-908-1
Perfect score: 23
Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 452

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	47.8	7	Q8KMS9	Q8kms9 enterobacte
2	11	47.8	7	Q8GL04	Q8gl04 borrelia bu
3	11	47.8	7	Q8GL00	Q8gl00 borrelia bu
4	11	47.8	7	Q85945	Q85945 saccharomyc
5	11	47.8	7	10 Q49223	Q49223 glycine max
6	11	47.8	7	15 Q8JEB1	Q8jeb1 human immun
7	11	47.8	8	2 Q85406	Q85406 coxiella bu
8	11	47.8	8	2 Q8GL21	Q8gl21 borrelia bu
9	11	47.8	8	2 Q8G940	Q8g940 borrelia bu
10	11	47.8	8	4 Q15890	Q15890 homo sapien
11	11	47.8	8	4 Q15888	Q15888 homo sapien
12	11	47.8	8	4 Q8IVK3	Q8ivk3 homo sapien
13	11	47.8	8	5 P82685	P82685 periplaneta
14	11	47.8	8	5 P82686	P82686 periplaneta
15	11	47.8	8	5 P82687	P82687 periplaneta
16	11	47.8	8	5 P82688	P82688 periplaneta

17	11	47.8	8	5 P82689	P82689 periplaneta
18	11	47.8	8	6 Q02831	Q02831 cryptolagus
19	11	47.8	8	6 Q95M23	Q95m23 sus scrofa
20	11	47.8	8	6 Q9TRY3	Q9try3 sus sp. ins
21	11	47.8	8	8 Q8W8G2	Q8w8g2 diadema sav
22	11	47.8	8	8 Q8W8G4	Q8w8g4 diadema mex
23	11	47.8	8	8 Q94VC1	Q94vc1 varanus rud
24	11	47.8	8	8 Q94VF6	Q94vf6 varanus job
25	11	47.8	8	8 Q8NGD7	Q8ngd7 louis hirta
26	11	47.8	8	8 Q94V88	Q94v88 varanus tri
27	11	47.8	8	8 Q94V82	Q94v82 varanus yuw
28	11	47.8	8	8 Q9TD02	Q9td02 terranatos
29	11	47.8	8	8 Q94Y2	Q94y2 asterina pe
30	11	47.8	8	8 Q94PX5	Q94px5 felis silve
31	11	47.8	8	8 Q8W8G6	Q8w8g6 diadema mex
32	11	47.8	8	8 Q8W8G5	Q8w8g5 diadema ant
33	11	47.8	8	8 Q94VJ4	Q94vj4 varanus ben
34	11	47.8	8	8 Q8W8G3	Q8w8g3 diadema pau
35	11	47.8	8	8 Q94V91	Q94v91 varanus tim
36	11	47.8	8	8 Q94VE4	Q94ve4 varanus mel
37	11	47.8	8	8 Q94VB2	Q94vb2 varanus sal
38	11	47.8	8	8 Q94VF9	Q94vf9 varanus ind
39	11	47.8	8	8 Q94PX7	Q94px7 felis silve
40	11	47.8	8	8 Q94PX6	Q94px6 felis libyc
41	11	47.8	8	8 Q94VA7	Q94va7 varanus sal
42	11	47.8	8	8 Q94VB5	Q94vb5 varanus sal
43	11	47.8	8	11 Q9ETL8	Q9etl8 mus spretus
44	11	47.8	8	11 Q99MNO	Q99mn0 mus musculu
45	11	47.8	8	11 Q35835	Q35835 rattus sp.

ALIGNMENTS

RESULT 1

Q8KMS9 ID Q8KMS9 PRELIMINARY; PRT; 7 AA.
AC Q8KMS9;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Putative transposase (Fragment).
GN TNIA.
OS Enterobacter sp. CH2-4.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Enterobacter.
OX NCBI_TaxID=143777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CH2-4;
RX MEDLINE=21604134; PubMed=11763242;
RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
RA Yurieva O.V., Nikiforov V.G.;
RT "Mercury resistance transposons of Gram-negative environmental
bacteria and their classification.";
RL Res. Microbiol. 152:811-822(2001).
DR EMBL; AJ302778; CAC83058.1; -;
FT NON_TER 1
SQ SEQUENCE 7 AA; 966 MW; 737B1B1046DAA9A0 CRC64;

Query Match 47.8%; Score 11; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1

DB 7 W 7

RESULT 2

Q8GL04 ID Q8GL04 PRELIMINARY; PRT; 7 AA.

AC Q8GL04;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE PF-50 protein (Fragment).
 CN PP-50.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid group cp32-5.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 ON NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=93-0107;
 RA Stevenson B., Miller J.C.;
 RT "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
 RT prophages: conservation amidst diversity.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY142103; AAN17848.1; -
 KW Plasmid.
 FT NON_TER
 SQ SEQUENCE 7 AA; 914 MW; 6337244330504310 CRC64;
 Query Match 47.8%; Score 11; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 W 1
 Db 1 W 1
 RESULT 3
 ID Q8GL00 PRELIMINARY; PRT; 7 AA.
 AC Q8GL00;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE PF-50 protein (Fragment).
 GN PF-50.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid group cp32-13.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 ON NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA15;
 RA Stevenson B., Miller J.C.;
 RT "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
 RT prophages: conservation amidst diversity.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY142106; AAN17857.1; -
 KW Plasmid.
 FT NON_TER
 SQ SEQUENCE 7 AA; 928 MW; 6337233050437350 CRC64;
 Query Match 47.8%; Score 11; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 W 1
 Db 2 W 2
 RESULT 4
 ID Q95945 PRELIMINARY; PRT; 7 AA.
 AC Q95945;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Inside intron 5 (Fragment).

OS Saccharomyces cerevisiae (Baker's yeast).
 OC Mitochondrion.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 ON NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D273-10B;
 RX MEDLINE=81069885; PubMed=6254986;
 RA Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;
 RT "Assembly of the mitochondrial membrane system: Structure and
 RT nucleotide sequence of the gene coding for subunit 1 of yeast
 RT cytochrome oxidase.";
 RL J. Biol. Chem. 255:11927-11941(1980).
 DR EMBL; V00694; CAA24066.1; -
 KW Mitochondrion.
 FT NON_TER
 SQ SEQUENCE 7 AA; 859 MW; 75B7232362CDC460 CRC64;
 Query Match 47.8%; Score 11; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 W 1
 Db 4 W 4
 RESULT 5
 ID 049223 PRELIMINARY; PRT; 7 AA.
 AC 049223;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE HMG-I-like protein (Fragment).
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 ON NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Essex; TISSUE=Root;
 RX MEDLINE=91367679; PubMed=1891369;
 RA Laux T., Goldberg R.B.;
 RT "A plant DNA binding protein shares highly conserved sequence motifs
 RT with HMG-box proteins.";
 RL Nucleic Acids Res. 19:4769-4769(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Essex; TISSUE=Root;
 RA Mahalingam R., Knap H.T.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF047050; AAC03556.1; -
 FT NON_TER
 SQ SEQUENCE 7 AA; 850 MW; 6AAAAAB378637810 CRC64;
 Query Match 47.8%; Score 11; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 W 1
 Db 2 W 2
 RESULT 6
 ID Q8JE81 PRELIMINARY; PRT; 7 AA.
 AC Q8JE81;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Truncated pol protein (Fragment).
 GN POL.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=4874;
 RX MEDLINE=22056123; PubMed=12060770;
 RA Beerewinkel N., Schmidt B., Walter H., Kaiser R., Lengauer T.,
 RA Hofmann D., Korn K., Selbig J.;
 RT "Diversity and complexity of HIV-1 drug resistance: A bioinformatics
 RT approach to predicting phenotype from genotype."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:8271-8276(2002).
 DR EMBL; AF347267; AK33344.1; -.
 FT NON_TER 1 1
 SQ SEQUENCE 7 AA; 885 MW; 76C37731A046C700 CRC64;

 Query Match 47.8%; Score 11; DB 15; Length 7;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 W 1
 Db 6 W 6

 RESULT 7
 O85406 PRELIMINARY; PRT; 8 AA.
 ID O85406;
 AC O85406;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 1.0 kDa protein (fragment).
 OS Coxiella burnetii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
 OC Coxiellaceae; Coxiella.
 OX NCBI_TaxID=777;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nine Mile Phase I;
 RX MEDLINE=98348442; PubMed=9683477;
 RA Williams H., Jaeger C., Baljer G.;
 RT "Physical and genetic map of the obligate intracellular bacterium
 RT Coxiella burnetii."
 RL J. Bacteriol. 180:3816-3822(1998).
 DR EMBL; AF064963; AAD09947.1; -.
 KW Hypothetical protein.
 FT NON_TER 1 1
 SQ SEQUENCE 8 AA; 993 MW; 04B55AA453772727 CRC64;

 Query Match 47.8%; Score 11; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 W 1
 Db 4 W 4

 RESULT 8
 O8GL21 PRELIMINARY; PRT; 8 AA.
 ID O8GL21
 AC O8GL21;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE PF-50 protein (fragment).
 GN PF-50.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid group cp32-8.

OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sh-2-82;
 RA Stevenson B., Miller J.C.;
 RT "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
 RT prophages: conservation amidst diversity."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY142094; AAN17903.1; -.
 KW Plasmid.
 FT NON_TER 1 1
 SQ SEQUENCE 8 AA; 1042 MW; 1437244337204373 CRC64;

 Query Match 47.8%; Score 11; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 W 1
 Db 2 W 2

 RESULT 9
 O8G940 PRELIMINARY; PRT; 8 AA.
 ID O8G940;
 AC O8G940;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE PF-50 protein (fragment).
 GN PF-50.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OG Plasmid group cp32-9, and Plasmid group cp32-12.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DNI27c19-2, and Sh-2-82;
 RC PLASMID-group cp32-9, and group cp32-12;
 RA Stevenson B., Miller J.C.;
 RT "Comparative analyses of Borrelia burgdorferi erp genes and their cp32
 RT prophages: conservation amidst diversity."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY142104; AAN17853.1; -.
 DR EMBL; AY142097; AAN17907.1; -.
 KW Plasmid.
 FT NON_TER 1 1
 SQ SEQUENCE 8 AA; 1042 MW; 1437244330504373 CRC64;

 Query Match 47.8%; Score 11; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 W 1
 Db 2 W 2

 RESULT 10
 Q15890 PRELIMINARY; PRT; 8 AA.
 ID Q15890;
 AC Q15890;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE (Clone XP19G12A) (fragment).
 OS Homo sapiens (Human)
 CC Zukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32083; AAA73880.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 975 MW; 605EA6C5BEA5A2D3 CRC64;

Query Match 47.8%; Score 11; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 1 W 1

RESULT 11
Q15888
ID Q15888 PRELIMINARY; PRT; 8 AA.
AC Q15888;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (Clone XP1588A) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL; L32069; AAA73878.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 1068 MW; 0315A37EAB5B0763 CRC64;

Query Match 47.8%; Score 11; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 6 W 6

RESULT 12
Q81VK3
ID Q81VK3 PRELIMINARY; PRT; 8 AA.
AC Q81VK3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Steerin2 (fragment).
GN STEERIN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Peeters P.J., Verhasselt P., Moechars D.W., Luyten W.H.M.L.,
RA Geysen J.J.G.H.;

```

```

RT "Sensorineural defects in mice hypomorphic for a mammalian homolog of
RT unc-53.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ488208; CAD32561.1; -.
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 1030 MW; E8A7245417737726 CRC64;

Query Match 47.8%; Score 11; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 3 W 3

RESULT 13
P82685
ID P82685 PRELIMINARY; PRT; 8 AA.
AC P82685;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Kinin-1 (PEA-K-1).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;
OC Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the
RT retrocerebral complex of the American cockroach, Periplaneta
RT americana.";
RL Regul. Pept. 71:199-205(1997).
CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC -!- FUNCTION: BELONGS TO THE KININ FAMILY.
CC -!- SIMILARITY: BELONGS TO THE KININ FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8
FT MOD_RES 8 8
SQ SEQUENCE 8 AA; 950 MW; 326365B449D5A774 CRC64;

Query Match 47.8%; Score 11; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 8.3e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
DB 7 W 7

RESULT 14
P82686
ID P82686 PRELIMINARY; PRT; 8 AA.
AC P82686;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Kinin-2 (PEA-K-2).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;
OC Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE=98010462; PubMed=9350979;
RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of eight kinins from the

```

RT retrocerebral complex of the American cockroach, Periplaneta
 americana.";
 RL Regul. Pept. 71:199-205(1997).
 CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
 (MYOTROPIC ACTIVITY)
 CC -!- SIMILARITY: BELONGS TO THE KININ FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD_RES 8 8
 SQ SEQUENCE 8 AA; 856 MW; DC6365A5B9D5BDDA CRC64;

Query Match 47.8%; Score 11; DB 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
 |
 Db 7 W 7

RESULT 15

P82687 PRELIMINARY; PRT; 8 AA.
 AC P82687
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DF 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE Kinin-3 (PEA-K-3).
 OS Periplaneta americana (American cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;
 OC Blattellidae; Periplaneta.
 OX NCBI_TaxID=6978;
 RN [1]
 RP SEQUENCE AND FUNCTION.
 RC TISSUE-CORPORA CARDIACA;
 RX MEDLINE=98010462; PubMed=9350979;
 RA Predel R., Kellner R., Rapus J., Penzlin H., Gaede G.;
 RT "Isolation and structural elucidation of eight kinins from the
 retrocerebral complex of the American cockroach, Periplaneta
 americana.";
 RL Regul. Pept. 71:199-205(1997).
 CC -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
 (MYOTROPIC ACTIVITY).
 CC -!- SIMILARITY: BELONGS TO THE KININ FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD_RES 8 8
 SQ SEQUENCE 8 AA; 909 MW; DC6365B449D5A76A CRC64;

Query Match 47.8%; Score 11; DB 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
 |
 Db 7 W 7

Search completed: August 16, 2003, 14:47:40
 Job time : 99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:35:34 : Search time 81 Seconds
(without alignments)
15.677 Million cell updates/sec

Title: US-09-185-908-1

Perfect score: 23

Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 102276

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

Result No.	Score	Query Match	Length	DB ID	Description
1	14	60.9	8	21	AAB06521
2	14	60.9	8	21	AAB06574
3	14	60.9	8	21	AAB06823
4	14	60.9	8	21	AAB06876
5	14	60.9	8	21	AAB06917
6	13	56.5	8	21	AAB56267
7	13	56.5	8	21	AAB06419
8	13	56.5	8	21	AAB06426
9	13	56.5	8	21	AAB06479
SUMMARIES					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
<p>1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*</p> <p>2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*</p> <p>3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*</p> <p>4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*</p> <p>5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*</p> <p>6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*</p> <p>7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*</p> <p>8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*</p> <p>9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*</p> <p>10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*</p> <p>11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*</p> <p>12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*</p> <p>13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*</p> <p>14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*</p> <p>15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*</p> <p>16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*</p> <p>17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*</p> <p>18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*</p> <p>19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*</p> <p>20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*</p> <p>21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*</p> <p>22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*</p> <p>23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*</p> <p>24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*</p>					

10	13	56.5	8	21	AAB06512	Claudin-1 cyclic c
11	13	56.5	8	21	AAB06583	Claudin-3 cell adh
12	13	56.5	8	21	AAB06636	Claudin-3 cyclic c
13	13	56.5	8	21	AAB06645	Claudin-4 cell adh
14	13	56.5	8	21	AAB06698	Claudin-4 cyclic c
15	13	56.5	8	21	AAB06764	Claudin-6/9 cell a
16	13	56.5	8	21	AAB06814	Claudin-6/9 cyclic
17	13	56.5	8	21	AAB06916	Claudin cell adhes
18	13	56.5	8	21	AAV77682	Hyaluronic acid (H
19	13	56.5	8	22	AAB51087	Dopamine transport
20	13	56.5	8	23	ABB1045	Tab fragment direc
21	12	52.2	8	20	AAV19720	SEQ ID NO 438 from
22	12	52.2	8	21	AAB09437	Hepatitis GB virus
23	12	52.2	8	23	ABJ10684	Japanese encephali
24	11	47.8	2	22	AAV97814	Human peptide #108
25	11	47.8	2	22	AAV98746	Human peptide #202
26	11	47.8	2	22	AAB91404	Tachykinins peptid
27	11	47.8	2	22	AAB91562	Endothelins and re
28	11	47.8	2	22	AAB91731	Oploid peptide SEQ
29	11	47.8	2	23	ABG63385	Human albumin fusi
30	11	47.8	2	24	ABR48109	Human secreted pro
31	11	47.8	3	10	AAV90982	Binding receptor w
32	11	47.8	3	10	AAV97798	Sequence of fragme
33	11	47.8	3	11	AAV05915	Partial sequence o
34	11	47.8	3	15	AAV65949	Tripeptide inhibit
35	11	47.8	3	15	AAV42556	ACE inhibitor SP3.
36	11	47.8	3	15	AAV47530	GRP-6 analog. Sy
37	11	47.8	3	15	AAV48519	Lactoferrin derive
38	11	47.8	3	15	AAV48523	Lactoferrin derive
39	11	47.8	3	15	AAV48527	Lactoferrin derive
40	11	47.8	3	15	AAV49510	Endothelin antagon
41	11	47.8	3	15	AAV57450	Lactoferrin derive
42	11	47.8	3	15	AAV57454	Lactoferrin derive
43	11	47.8	3	15	AAV57458	Lactoferrin derive
44	11	47.8	3	15	AAV58576	Angiotensin I conv
45	11	47.8	3	15	AAV58579	Angiotensin I conv

ALIGNMENTS

RESULT 1

AB06521	ID	AA06521	standard; peptide; 8 AA.
XX	AC	AA06521;	
XX	DT	28-SEP-2000	(first entry)
XX	DE	Claudin-2	cell adhesion recognition sequence SEQ ID NO: 42.
XX	KW	Claudin-2	modulating agent; cell adhesion recognition sequence;
XX	KW	CAR sequence;	autoimmune disease; inflammatory disease; cancer;
XX	KW	graft rejection.	
XX	OS	Mammalia.	
XX	PN	WO200026360-A1.	
XX	PD	11-MAY-2000.	
XX	PF	03-NOV-1999;	99WO-CA01029.
XX	PR	03-NOV-1998;	98US-0185908.
XX	PR	30-MAR-1999;	99US-0282029.
XX	PA	(ADHE-)	ADHEREX TECHNOLOGIES INC.
XX	PI	Blaschuck OW,	Symonds JM, Gour BJ;
XX	DR	WPI;	2000-365610/31.
XX	PT	Antibody modulation of claudin-mediated cell adhesion for increasing	

PT vasopermeability, for delivering drugs to tumors and the nervous system
PT and across the skin -
XX PS Claim 46; Page 97; 121pp; English.
XX CC The present invention relates to the use of peptides as claudin-mediated
CC cell adhesion modulators. The claudin-2 group of proteins are cadherins,
CC which are membrane glycoproteins involved in cell adhesion. In some
CC situations, cell adhesion occurs at abnormal levels, and these peptides
CC can be used to modulate these levels, and thus treat autoimmune diseases,
CC inflammatory diseases and cancer, and aid wound healing and implant
CC adhesion. In addition, they can also be used to facilitate drug delivery
CC to the desired target site.
XX SQ Sequence 8 AA;
XX
XX Query Match 60.9%; Score 14; DB 21; Length 8;
XX Best Local Similarity 25.0%; Pred. No. 9.3e+05;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 1 WXXXXXXG 8
XX | | | | |
XX Db 1 WRTSSVVG 8
XX
XX RESULT 2
XX ID AAB06574 standard; peptide; 8 AA.
XX AC AAB06574;
XX DT 28-SEP-2000 (first entry)
XX DE Claudin-2 cyclic cell adhesion recognition sequence SEQ ID NO: 163.
XX KW Claudin-2 modulating agent; cell adhesion recognition sequence;
XX KR CAR sequence; autoimmune disease; inflammatory disease; cancer;
XX KW graft rejection; cyclic.
XX OS Mammalia.
XX PN WO200026360-A1.
XX XX
XX PD 11-MAY-2000.
XX PF 03-NOV-1999; 99WO-CA01029.
XX PR 03-NOV-1998; 98US-0185908.
XX PR 30-MAR-1999; 99US-0282029.
XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
XX PI Blaschuck OW, Symonds JM, Gour BJ;
XX DR WPI; 2000-365610/31.
XX
XX PT Antibody modulation of claudin-mediated cell adhesion for increasing
PT vasopermeability, for delivering drugs to tumors and the nervous system
PT and across the skin -
XX PS Claim 49; Page 98; 121pp; English.
XX CC The present invention relates to the use of peptides as claudin-mediated
CC cell adhesion modulators. The claudin-2 group of proteins are cadherins,
CC which are membrane glycoproteins involved in cell adhesion. In some
CC situations, cell adhesion occurs at abnormal levels, and these peptides
CC can be used to modulate these levels, and thus treat autoimmune diseases,
CC inflammatory diseases and cancer, and aid wound healing and implant
CC adhesion. In addition, they can also be used to facilitate drug delivery
CC to the desired target site. The present sequence has a cyclic
CC conformation.
XX SQ Sequence 8 AA;
XX
XX Query Match 60.9%; Score 14; DB 21; Length 8;
XX Best Local Similarity 25.0%; Pred. No. 9.3e+05;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 1 WXXXXXXG 8
XX | | | | |
XX Db 1 WQMSVYAG 8
XX
XX RESULT 4
XX ID AAB06876 standard; peptide; 8 AA.
XX AC AAB06876;
XX XX

Query Match 60.9%; Score 14; DB 21; Length 8;
Best Local Similarity 25.0%; Pred. No. 9.3e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
| | | | |
Db 1 WRTSSVVG 8

RESULT 3
ID AAB06823 standard; peptide; 8 AA.
XX AC AAB06823;
XX DT 28-SEP-2000 (first entry)
XX DE Claudin-7 cell adhesion recognition sequence SEQ ID NO: 394.
XX KW Claudin-7 modulating agent; cell adhesion recognition sequence;
XX KR CAR sequence; autoimmune disease; inflammatory disease; cancer;
XX KW graft rejection.
XX OS Mammalia.
XX PN WO200026360-A1.
XX XX
XX PD 11-MAY-2000.
XX PF 03-NOV-1999; 99WO-CA01029.
XX PR 03-NOV-1998; 98US-0185908.
XX PR 30-MAR-1999; 99US-0282029.
XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
XX PI Blaschuck OW, Symonds JM, Gour BJ;
XX DR WPI; 2000-365610/31.
XX
XX PT Antibody modulation of claudin-mediated cell adhesion for increasing
PT vasopermeability, for delivering drugs to tumors and the nervous system
PT and across the skin -
XX PS Claim 76; Page 104; 121pp; English.
XX CC The present invention relates to the use of peptides as claudin-mediated
CC cell adhesion modulators. The claudin-7 group of proteins are cadherins,
CC which are membrane glycoproteins involved in cell adhesion. In some
CC situations, cell adhesion occurs at abnormal levels, and these peptides
CC can be used to modulate these levels, and thus treat autoimmune diseases,
CC inflammatory diseases and cancer, and aid wound healing and implant
CC adhesion. In addition, they can also be used to facilitate drug delivery
CC to the desired target site.
XX SQ Sequence 8 AA;
XX
XX Query Match 60.9%; Score 14; DB 21; Length 8;
XX Best Local Similarity 25.0%; Pred. No. 9.3e+05;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 1 WXXXXXXG 8
XX | | | | |
XX Db 1 WQMSVYAG 8
XX
XX RESULT 4
XX ID AAB06876 standard; peptide; 8 AA.
XX AC AAB06876;
XX XX

XX PI Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;
 XX PI Duan RD, Florence KA, Soppet DR;
 XX DR WPI; 2000-579828/66.
 XX PT Isolated nucleic acid molecule encoding a human secreted protein is
 XX PT used in preventing, treating or ameliorating a medical condition -
 XX PS Disclosure; Page 1021; 1065pp; English.
 XX CC The polynucleotide sequences given in AAC99818 to AAC99977 encode the
 CC human secreted proteins given in AAB56077 to AAB56362. Human secreted
 CC proteins have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: immunosuppressive;
 CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
 CC vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;
 CC virucide; fungicide; and ophthalmological. The human secreted
 CC polynucleotides and proteins can be used to prevent, treat or ameliorate
 CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
 CC dogs, chickens or sheep. They are also used in diagnosing a pathological
 CC condition or susceptibility to a pathological condition. Disorders which
 CC are diagnosed or treated include autoimmune diseases e.g. rheumatoid
 CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
 CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
 CC disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders
 CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
 CC fungi and ocular disorders e.g. corneal infection. The proteins can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a
 CC food additive or preservative to increase or decrease storage
 CC capabilities. AAC99809 to AAC99817 and AAB56076 represent sequences used
 CC in the exemplification of the present invention.
 XX SQ Sequence 8 AA;
 Query Match 56.5%; Score 13; DB 21; Length 8;
 Best Local Similarity 25.0%; Pred. No. 9.3e-05;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 WXXXXXXG 8
 DB 1 WGLVTLAG 8
 RESULT 7
 AAB06419
 ID AAB06419 standard; peptide; 8 AA.
 XX AC AAB06419;
 XX DT 28-SEP-2000 (first entry)
 XX DE Claudin-1 cell adhesion recognition sequence SEQ ID NO: 34.
 XX KW Claudin-1 modulating agent; cell adhesion recognition sequence;
 KW CAR sequence; autoimmune disease; inflammatory disease; cancer;
 KW graft rejection.
 XX OS Mammalia.
 XX PN WO200026360-A1.
 XX PD 11-MAY-2000.
 XX PF 03-NOV-1999; 99WO-CA01029.
 XX PR 03-NOV-1998; 98US-0185908.
 XX PR 30-MAR-1999; 99US-0282029.
 XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX PI Blaschuck OW, Symonds JM, Gour BJ;
 XX DR WPI; 2000-365610/31.
 XX PT Antibody modulation of claudin-mediated cell adhesion for increasing
 XX PT vasopermeability, for delivering drugs to tumors and the nervous system
 XX PT and across the skin -
 XX PS Claim 39; Page 96; 121pp; English.
 XX CC The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-1 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,
 CC inflammatory diseases and cancer, and aid wound healing and implant
 CC adhesion. In addition, they can also be used to facilitate drug delivery
 CC to the desired target site.
 XX SQ Sequence 8 AA;
 Query Match 56.5%; Score 13; DB 21; Length 8;
 Best Local Similarity 25.0%; Pred. No. 9.3e-05;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 WXXXXXXG 8
 DB 1 WKIYSYAG 8
 RESULT 8
 AAB06426
 ID AAB06426 standard; peptide; 8 AA.
 XX AC AAB06426;
 XX DT 28-SEP-2000 (first entry)
 XX DE Claudin-1 cell adhesion recognition sequence SEQ ID NO: 485.
 XX KW Claudin-1 modulating agent; cell adhesion recognition sequence;
 KW CAR sequence; autoimmune disease; inflammatory disease; cancer;
 KW graft rejection.
 XX OS Mammalia.
 XX PN WO200026360-A1.
 XX PD 11-MAY-2000.
 XX PF 03-NOV-1999; 99WO-CA01029.
 XX PR 03-NOV-1998; 98US-0185908.
 XX PR 30-MAR-1999; 99US-0282029.
 XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX PI Blaschuck OW, Symonds JM, Gour BJ;
 XX DR WPI; 2000-365610/31.
 XX PT Antibody modulation of claudin-mediated cell adhesion for increasing
 XX PT vasopermeability, for delivering drugs to tumors and the nervous system
 XX PT and across the skin -
 XX PS Claim 39; Page 96; 121pp; English.
 XX CC The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-1 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,

XX PI Blaschuck OW, Symonds JM, Gour BJ;
 XX DR WPI; 2000-365610/31.
 XX PT Antibody modulation of claudin-mediated cell adhesion for increasing
 XX PT vasopermeability, for delivering drugs to tumors and the nervous system
 XX PT and across the skin -
 XX PS Claim 39; Page 96; 121pp; English.
 XX CC The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-1 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,
 CC inflammatory diseases and cancer, and aid wound healing and implant
 CC adhesion. In addition, they can also be used to facilitate drug delivery
 CC to the desired target site.
 XX SQ Sequence 8 AA;
 Query Match 56.5%; Score 13; DB 21; Length 8;
 Best Local Similarity 25.0%; Pred. No. 9.3e-05;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 WXXXXXXG 8
 DB 1 WKIYSYAG 8
 RESULT 8
 AAB06426
 ID AAB06426 standard; peptide; 8 AA.
 XX AC AAB06426;
 XX DT 28-SEP-2000 (first entry)
 XX DE Claudin-1 cell adhesion recognition sequence SEQ ID NO: 485.
 XX KW Claudin-1 modulating agent; cell adhesion recognition sequence;
 KW CAR sequence; autoimmune disease; inflammatory disease; cancer;
 KW graft rejection.
 XX OS Mammalia.
 XX PN WO200026360-A1.
 XX PD 11-MAY-2000.
 XX PF 03-NOV-1999; 99WO-CA01029.
 XX PR 03-NOV-1998; 98US-0185908.
 XX PR 30-MAR-1999; 99US-0282029.
 XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX PI Blaschuck OW, Symonds JM, Gour BJ;
 XX DR WPI; 2000-365610/31.
 XX PT Antibody modulation of claudin-mediated cell adhesion for increasing
 XX PT vasopermeability, for delivering drugs to tumors and the nervous system
 XX PT and across the skin -
 XX PS Claim 39; Page 96; 121pp; English.
 XX CC The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-1 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,

CC inflammatory diseases and cancer, and aid wound healing and implant
 CC adhesion. In addition, they can also be used to facilitate drug delivery
 CC to the desired target site.

XX SQ Sequence 8 AA;
 Query Match 56.5%; Score 13; DB 21; Length 8;
 Best Local Similarity 25.0%; Pred. No. 9.3e+05;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 WXXXXXXG 8
 |
 Db 1 WKIYSYAG 8

RESULT 9
 AAB06479
 ID AAB06479 standard; peptide; 8 AA.
 XX AC AAB06479;
 XX DT 28-SEP-2000 (first entry)
 XX DE Claudin-1 cyclic cell adhesion recognition sequence SEQ ID NO: 110.
 XX KW Claudin-1 modulating agent; cell adhesion recognition sequence;
 KW CAR sequence; autoimmune disease; inflammatory disease; cancer;
 KW graft rejection; cyclic.
 XX OS Mammalia.

XX PN WO200026360-A1.
 XX PD 11-MAY-2000.
 XX PF 03-NOV-1999; 99WO-CA01029.
 XX PR 03-NOV-1998; 98US-0185908.
 XX PR 30-MAR-1999; 99US-0282029.
 XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX PI Blaschuck OW, Symonds JM, Gour BJ;
 XX DR WPI; 2000-365610/31.
 XX PT Antibody modulation of claudin-mediated cell adhesion for increasing
 PT vasopermeability, for delivering drugs to tumors and the nervous system
 PT and across the skin -
 XX PS Claim 43; Page 97; 121pp; English.

XX The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-1 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,
 CC inflammatory diseases and cancer, and aid wound healing and implant
 CC adhesion. In addition, they can also be used to facilitate drug delivery
 CC to the desired target site. The present sequence has a cyclic
 CC conformation.

XX SQ Sequence 8 AA;
 Query Match 56.5%; Score 13; DB 21; Length 8;
 Best Local Similarity 25.0%; Pred. No. 9.3e+05;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 WXXXXXXG 8
 |
 Db 1 WKIYSYAG 8

RESULT 10
 AAB06512
 ID AAB06512 standard; peptide; 8 AA.
 XX AC AAB06512;
 XX DT 28-SEP-2000 (first entry)
 XX DE Claudin-1 cyclic cell adhesion recognition sequence SEQ ID NO: 519.
 XX KW Claudin-1 modulating agent; cell adhesion recognition sequence;
 KW CAR sequence; autoimmune disease; inflammatory disease; cancer;
 KW graft rejection; cyclic.
 XX OS Mammalia.

XX PN WO200026360-A1.
 XX PD 11-MAY-2000.
 XX PF 03-NOV-1999; 99WO-CA01029.
 XX PR 03-NOV-1998; 98US-0185908.
 XX PR 30-MAR-1999; 99US-0282029.
 XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX PI Blaschuck OW, Symonds JM, Gour BJ;
 XX DR WPI; 2000-365610/31.
 XX PT Antibody modulation of claudin-mediated cell adhesion for increasing
 PT vasopermeability, for delivering drugs to tumors and the nervous system
 PT and across the skin -
 XX PS Claim 43; Page 97; 121pp; English.

XX The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-1 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,
 CC inflammatory diseases and cancer, and aid wound healing and implant
 CC adhesion. In addition, they can also be used to facilitate drug delivery
 CC to the desired target site. The present sequence has a cyclic
 CC conformation.

XX SQ Sequence 8 AA;
 Query Match 56.5%; Score 13; DB 21; Length 8;
 Best Local Similarity 25.0%; Pred. No. 9.3e+05;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 WXXXXXXG 8
 |
 Db 1 WKIYSYAG 8

RESULT 11
 AAB06583
 ID AAB06583 standard; peptide; 8 AA.
 XX AC AAB06583;
 XX DT 28-SEP-2000 (first entry)
 XX DE Claudin-3 cell adhesion recognition sequence SEQ ID NO: 58.
 XX KW Claudin-3 modulating agent; cell adhesion recognition sequence;
 KW CAR sequence; autoimmune disease; inflammatory disease; cancer;
 KW graft rejection.
 XX OS Mammalia.

XX WO2000026360-A1.
 XX 11-MAY-2000.
 XX 03-NOV-1999; 99WO-CA01029.
 XX 03-NOV-1998; 98US-0185908.
 XX 30-MAR-1999; 99US-0282029.
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX Blaschuck OW, Symonds JM, Gour BJ;
 XX WPI; 2000-365610/31.
 XX Antibody modulation of claudin-mediated cell adhesion for increasing
 PT vasopermeability, for delivering drugs to tumors and the nervous system
 PT and across the skin -
 XX Claim 52; Page 99; 121pp; English.
 XX The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-3 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,
 CC inflammatory diseases and cancer, and aid wound healing and implant
 CC adhesion. In addition, they can also be used to facilitate drug delivery
 CC to the desired target site.
 XX Sequence 8 AA;
 XX Query Match 56.5%; Score 13; DB 21; Length 8;
 XX Best Local Similarity 25.0%; Pred. No. 9.3e+05;
 XX Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 OY 1 WXXXXXXG 8
 DB 1 WRVSAFIG 8
 RESULT 12
 AAB06636
 ID AAB06636 standard; peptide; 8 AA.
 XX AC AAB06636;
 XX 28-SEP-2000 (first entry)
 XX Claudin-3 cyclic cell adhesion recognition sequence SEQ ID NO: 269.
 XX Claudin-3 modulating agent; cell adhesion recognition sequence;
 KW CAR sequence; autoimmune disease; inflammatory disease; cancer;
 KW graft rejection; cyclic.
 XX Mammalia.
 XX WO2000026360-A1.
 XX 11-MAY-2000.
 XX 03-NOV-1999; 99WO-CA01029.
 XX 03-NOV-1998; 98US-0185908.
 XX 30-MAR-1999; 99US-0282029.
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX Blaschuck OW, Symonds JM, Gour BJ;
 XX WPI; 2000-365610/31.
 XX Antibody modulation of claudin-mediated cell adhesion for increasing
 PT vasopermeability, for delivering drugs to tumors and the nervous system
 PT and across the skin -
 XX Claim 52; Page 99; 121pp; English.
 XX The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-3 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,
 CC inflammatory diseases and cancer, and aid wound healing and implant
 CC adhesion. In addition, they can also be used to facilitate drug delivery
 CC to the desired target site.
 XX Sequence 8 AA;
 XX Query Match 56.5%; Score 13; DB 21; Length 8;
 XX Best Local Similarity 25.0%; Pred. No. 9.3e+05;
 XX Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 OY 1 WXXXXXXG 8
 DB 1 WRVSAFIG 8

PT Antibody modulation of claudin-mediated cell adhesion for increasing
 PT vasopermeability, for delivering drugs to tumors and the nervous system
 PT and across the skin -
 XX Claim 55; Page 100; 121pp; English.
 XX The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-3 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,
 CC inflammatory diseases and cancer, and aid wound healing and implant
 CC adhesion. In addition, they can also be used to facilitate drug delivery
 CC to the desired target site. The present sequence has a cyclic
 CC conformation.
 XX Sequence 8 AA;
 XX Query Match 56.5%; Score 13; DB 21; Length 8;
 XX Best Local Similarity 25.0%; Pred. No. 9.3e+05;
 XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 OY 1 WXXXXXXG 8
 DB 1 WRVSAFIG 8
 RESULT 13
 AAB06645
 ID AAB06645 standard; peptide; 8 AA.
 XX AC AAB06645;
 XX 28-SEP-2000 (first entry)
 XX Claudin-4 cell adhesion recognition sequence SEQ ID NO: 50.
 XX Claudin-4 modulating agent; cell adhesion recognition sequence;
 KW CAR sequence; autoimmune disease; inflammatory disease; cancer;
 KW graft rejection.
 XX Mammalia.
 XX WO2000026360-A1.
 XX 11-MAY-2000.
 XX 03-NOV-1999; 99WO-CA01029.
 XX 03-NOV-1998; 98US-0185908.
 XX 30-MAR-1999; 99US-0282029.
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX Blaschuck OW, Symonds JM, Gour BJ;
 XX WPI; 2000-365610/31.
 XX Antibody modulation of claudin-mediated cell adhesion for increasing
 PT vasopermeability, for delivering drugs to tumors and the nervous system
 PT and across the skin -
 XX Claim 58; Page 100; 121pp; English.
 XX The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-4 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,
 CC inflammatory diseases and cancer, and aid wound healing and implant
 CC adhesion. In addition, they can also be used to facilitate drug delivery
 CC to the desired target site.

SQ Sequence 8 AA;

Query Match 56.5%; Score 13; DB 21; Length 8;
 Best Local Similarity 25.0%; Pred. No. 9.3e+05;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 | | | | |
 Db 1 WRVTAFIG 8

RESULT 14

AAB06698
 ID AAB06698 standard; peptide; 8 AA.

AC AAB06698;

DT 28-SEP-2000 (first entry)

DE Claudin-4 cyclic cell adhesion recognition sequence SEQ ID NO: 216.

KW Claudin-4 modulating agent; cell adhesion recognition sequence;
 KW CAR sequence; autoimmune disease; inflammatory disease; cancer;
 KW graft rejection; cyclic.

XX Mammalia.

XX WO200026360-AL.

XX 11-MAY-2000.

XX 03-NOV-1999; 99WO-CA01029.

XX 03-NOV-1998; 98US-0185908.

XX 30-MAR-1999; 99US-0282029.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuck OW, Symonds JM, Gour BJ;

XX WPI; 2000-365610/31.

PT Antibody modulation of claudin-mediated cell adhesion for increasing
 PT vasopermeability, for delivering drugs to tumors and the nervous system
 PT and across the skin -

PS Claim 61; Page 101; 121pp; English.

XX The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-4 group of proteins are cadherins,
 CC which are membrane glycoproteins involved in cell adhesion. In some
 CC situations, cell adhesion occurs at abnormal levels, and these peptides
 CC can be used to modulate these levels, and thus treat autoimmune diseases,
 CC inflammatory diseases and cancer, and aid wound healing and implant
 CC adhesion. In addition, they can also be used to facilitate drug delivery
 CC to the desired target site. The present sequence has a cyclic
 CC conformation.

SQ Sequence 8 AA;

Query Match 56.5%; Score 13; DB 21; Length 8;
 Best Local Similarity 25.0%; Pred. No. 9.3e+05;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 | | | | |
 Db 1 WRVTAFIG 8

RESULT 15

AAB06764
 ID AAB06764 standard; peptide; 8 AA.

XX

AC AAB06764;

XX 28-SEP-2000 (first entry)

DE Claudin-6/9 cell adhesion recognition sequence SEQ ID NO: 335.

KW Claudin-6 modulating agent; claudin-9 modulating agent;
 KW cell adhesion recognition sequence; CAR sequence; autoimmune disease;
 KW inflammatory disease; cancer; graft rejection.

XX Mammalia.

XX WO200026360-AL.

XX 11-MAY-2000.

XX 03-NOV-1999; 99WO-CA01029.

XX 03-NOV-1998; 98US-0185908.

XX 30-MAR-1999; 99US-0282029.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuck OW, Symonds JM, Gour BJ;

XX WPI; 2000-365610/31.

PT Antibody modulation of claudin-mediated cell adhesion for increasing
 PT vasopermeability, for delivering drugs to tumors and the nervous system
 PT and across the skin -

PS Claim 70; Page 103; 121pp; English.

XX The present invention relates to the use of peptides as claudin-mediated
 CC cell adhesion modulators. The claudin-6 and claudin-9 groups of proteins
 CC are cadherins, which are membrane glycoproteins involved in cell
 CC adhesion. In some situations, cell adhesion occurs at abnormal levels,
 CC and these peptides can be used to modulate these levels, and thus treat
 CC autoimmune diseases, inflammatory diseases and cancer, and aid wound
 CC healing and implant adhesion. In addition, they can also be used to
 CC facilitate drug delivery to the desired target site.

SQ Sequence 8 AA;

Query Match 56.5%; Score 13; DB 21; Length 8;
 Best Local Similarity 25.0%; Pred. No. 9.3e+05;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 | | | | |
 Db 1 WKVTAFIG 8

Search completed: August 16, 2003, 14:45:26
 Job time : 82 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:47:45 ; Search time 51 seconds
(without alignments)
20.550 Million cell updates/sec

Title: US-09-185-908-1
Perfect score: 23
Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 492763 seqs, 131003257 residues

Total number of hits satisfying chosen parameters: 35679

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	60.9	8	10	US-09-185-908-42
2	14	60.9	8	10	US-09-185-908-163
3	13	56.5	8	10	US-09-185-908-34
4	13	56.5	8	10	US-09-185-908-50
5	13	56.5	8	10	US-09-185-908-58
6	13	56.5	8	10	US-09-185-908-110
7	13	56.5	8	10	US-09-185-908-216
8	13	56.5	8	10	US-09-185-908-269
9	13	56.5	8	11	US-09-986-480-361
10	13	56.5	8	15	US-10-119-537-11
11	12	52.2	8	8	US-08-424-550B-564
12	11	47.8	3	2	US-10-197-954-42
13	11	47.8	3	10	US-09-922-261-291
14	11	47.8	3	10	US-09-922-261-405
15	11	47.8	3	10	US-09-982-172-25

16	11	47.8	3	11	US-09-809-391-525	Sequence 525, App
17	11	47.8	3	11	US-09-852-910-151	Sequence 151, App
18	11	47.8	3	12	US-10-130-082-239	Sequence 239, App
19	11	47.8	3	15	US-10-039-876A-6	Sequence 6, Appli
20	11	47.8	4	8	US-08-484-409-5	Sequence 5, Appli
21	11	47.8	4	8	US-08-484-409-29	Sequence 29, Appl
22	11	47.8	4	8	US-08-484-409-34	Sequence 34, Appl
23	11	47.8	4	8	US-08-981-824-43	Sequence 43, Appl
24	11	47.8	4	8	US-08-424-550B-216	Sequence 216, App
25	11	47.8	4	8	US-08-424-550B-314	Sequence 314, App
26	11	47.8	4	8	US-08-424-550B-346	Sequence 346, App
27	11	47.8	4	8	US-08-841-636A-3	Sequence 3, Appli
28	11	47.8	4	9	US-09-736-611-1	Sequence 1, Appli
29	11	47.8	4	9	US-09-736-611-2	Sequence 2, Appli
30	11	47.8	4	9	US-09-765-527-196	Sequence 196, App
31	11	47.8	4	9	US-09-780-070-1	Sequence 1, Appli
32	11	47.8	4	9	US-09-780-070-2	Sequence 2, Appli
33	11	47.8	4	9	US-09-871-212-2	Sequence 2, Appli
34	11	47.8	4	9	US-09-808-037-4	Sequence 4, Appli
35	11	47.8	4	9	US-09-153-076-4	Sequence 4, Appli
36	11	47.8	4	9	US-09-153-076-5	Sequence 5, Appli
37	11	47.8	4	9	US-09-853-918-11	Sequence 11, Appl
38	11	47.8	4	9	US-09-853-918-19	Sequence 19, Appl
39	11	47.8	4	9	US-09-782-980-112	Sequence 112, App
40	11	47.8	4	9	US-09-741-148A-13	Sequence 13, Appl
41	11	47.8	4	9	US-09-682-667-12	Sequence 12, Appl
42	11	47.8	4	9	US-09-682-667-13	Sequence 13, Appl
43	11	47.8	4	9	US-09-746-742-23	Sequence 23, Appl
44	11	47.8	4	9	US-09-873-676-105	Sequence 104, App
45	11	47.8	4	9	US-09-873-676-104	Sequence 105, App

ALIGNMENTS

RESULT 1
US-09-185-908-42
; Sequence 42, Application US/09185908A
; Publication No. US20020193294A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; TITLE OF INVENTION: FUNCTIONS
; FILE REFERENCE: 100086.409
; CURRENT APPLICATION NUMBER: US/09/185,908A
; CURRENT FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on mouse claudin-2 sequence
US-09-185-908-42

Query Match 60.9%; Score 14; DB 10; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 1 WXXXXXXG 8
|
Db 1 WFTSSYVG 8

RESULT 2
US-09-185-908-163
; Sequence 163, Application US/09185908A
; Publication No. US20020193294A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.

```
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409
; CURRENT APPLICATION NUMBER: US/09/185,908A
; CURRENT FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 163
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on mouse claudin-2 sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide
US-09-185-908-163

Query Match          60.9%; Score 14; DB 10; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
   |         |
DB 1 WRTSSYVG 8

RESULT 3
US-09-185-908-34
; Sequence 34, Application US/09185908A
; Publication No. US20020193294A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409
; CURRENT APPLICATION NUMBER: US/09/185,908A
; CURRENT FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: Synthesis based on mouse claudin-1 sequence
US-09-185-908-34

Query Match          56.5%; Score 13; DB 10; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
   |         |
DB 1 WKIYSYAG 8

RESULT 4
US-09-185-908-50
; Sequence 50, Application US/09185908A
; Publication No. US20020193294A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409
; CURRENT APPLICATION NUMBER: US/09/185,908A
; CURRENT FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 269
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 50
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on human, mouse and monkey CPB-R
; OTHER INFORMATION: sequences
US-09-185-908-50

Query Match          56.5%; Score 13; DB 10; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
   |         |
DB 1 WRVTAFIG 8

RESULT 5
US-09-185-908-58
; Sequence 58, Application US/09185908A
; Publication No. US20020193294A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409
; CURRENT APPLICATION NUMBER: US/09/185,908A
; CURRENT FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on human and rat RVP-1 sequences
US-09-185-908-58

Query Match          56.5%; Score 13; DB 10; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
   |         |
DB 1 WRVSAFIG 8

RESULT 6
US-09-185-908-110
; Sequence 110, Application US/09185908A
; Publication No. US20020193294A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409
; CURRENT APPLICATION NUMBER: US/09/185,908A
; CURRENT FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 110
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on mouse claudin-1 sequence
```

;
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide
US-09-185-908-110

Query Match 56.5%; Score 13; DB 10; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 1 WKIYSYAG 8

RESULT 7

US-09-185-908-216
; Sequence 216, Application US/09185908A
; Publication No. US20020193294A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; TITLE OF INVENTION: FUNCTIONS
; FILE REFERENCE: 100086.409
; CURRENT APPLICATION NUMBER: US/09/185,908A
; CURRENT FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 216
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on human, mouse and monkey CPE-R
; OTHER INFORMATION: sequences
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide
US-09-185-908-216

Query Match 56.5%; Score 13; DB 10; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 1 WKIYSYAG 8

RESULT 8

US-09-185-908-269
; Sequence 269, Application US/09185908A
; Publication No. US20020193294A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; TITLE OF INVENTION: FUNCTIONS
; FILE REFERENCE: 100086.409
; CURRENT APPLICATION NUMBER: US/09/185,908A
; CURRENT FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 269
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on human and rat RVP-1 sequences
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide
US-09-185-908-269

Query Match 56.5%; Score 13; DB 10; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 1 WRVSAFIG 8

RESULT 9

US-09-986-480-361
; Sequence 361, Application US/09986480
; Publication No. US20030027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: PS500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 361
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-986-480-361

Query Match 56.5%; Score 13; DB 11; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 1 WGLVTLAG 8

RESULT 10

US-10-119-537-11
; Sequence 11, Application US/10119537
; Publication No. US20030027761A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: JUNCTIONAL ADHESION MOLECULE-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.412D1
; CURRENT APPLICATION NUMBER: US/10/119,537
; CURRENT FILING DATE: 2002-04-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fab fragment directed against claudin cell
; OTHER INFORMATION: adhesion recognition sequence
US-10-119-537-11

Query Match 56.5%; Score 13; DB 15; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 1 WKIYSYAG 8

RESULT 11

US-08-424-550B-564
; Sequence 564, Application US/08424550B
; Publication No. US2002011947A1
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSAHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550B
; FILING DATE:
; CLASSIFICATION: 435435
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 564:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-424-550B-564

Query Match 52.2%; Score 12; DB 8; Length 8;
Best Local Similarity 25.0%; Pred. No. 4.4e-05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXG 8
Db 1 WSGLLG 8

RESULT 12

US-10-197-954-42
; Sequence 42, Application US/10197954
; Publication No. US2003011902A1
; GENERAL INFORMATION:
; APPLICANT: K*ster, Hubert
; APPLICANT: Siddiqi, Suhail
; APPLICANT: Little, Daniel
; TITLE OF INVENTION: Capture Compounds, Collections Thereof
; TITLE OF INVENTION: And Methods For Analyzing The Proteome And Complex
; TITLE OF INVENTION: Compositions
; FILE REFERENCE: 24743-2305
; CURRENT APPLICATION NUMBER: US/10/197,954
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 60/306,019

; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/314,123
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/363,433
; PRIOR FILING DATE: 2002-03-11
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 2
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-197-954-42

Query Match 47.8%; Score 11; DB 15; Length 2;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
Db 1 W 1

RESULT 13

US-09-922-261-291
; Sequence 291, Application US/09922261
; Patent No. US20020111471A1
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puraanam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/922,261
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/09/461,697
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 291
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-261-291

Query Match 47.8%; Score 11; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
Db 2 W 2

RESULT 14

US-09-922-261-405
; Sequence 405, Application US/09922261
; Patent No. US20020111471A1
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puraanam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING

; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/922,261
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/09/461,697
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 405
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-922-261-405

Query Match 47.8%; Score 11; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 4.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 W 1
|
Db 2 W 2

RESULT 15

US-09-982-172-25
; Sequence 25, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-25

Query Match 47.8%; Score 11; DB 10; Length 3;
Best Local Similarity 100.0%; Pred. No. 4.4e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 W 1
|
Db 2 W 2

Search completed: August 16, 2003, 14:56:38
Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:45:34 ; Search time 366 Seconds
(without alignments)
19.024 Million cell updates/sec

Title: US-09-185-908-1

Perfect score: 23

Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 262382

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14	60.9	8	15	US-09-185-908-42
2	14	60.9	8	15	US-09-185-908-163

3	14	60.9	8	16	US-09-282-029-42	Sequence 42, Appl
4	14	60.9	8	16	US-09-282-029-163	Sequence 163, App
5	14	60.9	8	16	US-09-282-029-394	Sequence 394, App
6	14	60.9	8	16	US-09-282-029-447	Sequence 447, App
7	14	60.9	8	16	US-09-282-029A-42	Sequence 42, Appl
8	14	60.9	8	16	US-09-282-029A-163	Sequence 163, App
9	14	60.9	8	16	US-09-282-029A-394	Sequence 394, App
10	14	60.9	8	16	US-09-282-029A-447	Sequence 447, App
11	14	60.9	8	18	US-09-434-355-42	Sequence 42, Appl
12	14	60.9	8	18	US-09-434-355-163	Sequence 163, App
13	14	60.9	8	18	US-09-434-355-394	Sequence 394, App
14	14	60.9	8	18	US-09-434-355-447	Sequence 447, App
15	14	60.9	8	18	US-09-434-355A-42	Sequence 42, Appl
16	14	60.9	8	18	US-09-434-355A-163	Sequence 163, App
17	14	60.9	8	18	US-09-434-355A-394	Sequence 394, App
18	14	60.9	8	18	US-09-434-355A-447	Sequence 447, App
19	13	56.5	8	1	PCT-US00-12788-361	Sequence 361, App
20	13	56.5	8	15	US-09-185-908-34	Sequence 34, Appl
21	13	56.5	8	15	US-09-185-908-58	Sequence 58, Appl
22	13	56.5	8	15	US-09-185-908-110	Sequence 110, App
23	13	56.5	8	15	US-09-185-908-216	Sequence 216, App
24	13	56.5	8	15	US-09-185-908-269	Sequence 269, App
25	13	56.5	8	16	US-09-282-029-34	Sequence 34, Appl
26	13	56.5	8	16	US-09-282-029-50	Sequence 50, Appl
27	13	56.5	8	16	US-09-282-029-58	Sequence 58, Appl
28	13	56.5	8	16	US-09-282-029-110	Sequence 110, App
29	13	56.5	8	16	US-09-282-029-216	Sequence 216, App
30	13	56.5	8	16	US-09-282-029-269	Sequence 269, App
31	13	56.5	8	16	US-09-282-029-335	Sequence 335, App
32	13	56.5	8	16	US-09-282-029-385	Sequence 385, App
33	13	56.5	8	16	US-09-282-029A-34	Sequence 34, Appl
34	13	56.5	8	16	US-09-282-029A-50	Sequence 50, Appl
35	13	56.5	8	16	US-09-282-029A-58	Sequence 58, Appl
36	13	56.5	8	16	US-09-282-029A-110	Sequence 110, App
37	13	56.5	8	16	US-09-282-029A-216	Sequence 216, App
38	13	56.5	8	16	US-09-282-029A-269	Sequence 269, App
39	13	56.5	8	16	US-09-282-029A-335	Sequence 335, App
40	13	56.5	8	16	US-09-282-029A-385	Sequence 385, App
41	13	56.5	8	17	US-09-347-707A-15	Sequence 15, Appl
42	13	56.5	8	17	US-09-347-707B-15	Sequence 15, Appl
43	13	56.5	8	18	US-09-434-355-34	Sequence 34, Appl
44	13	56.5	8	18	US-09-434-355-50	Sequence 50, Appl
45	13	56.5	8	18		

ALIGNMENTS

RESULT 1
US-09-185-908-42
; Sequence 42, Application US/09185908A
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; TITLE OF INVENTION: FUNCTIONS
; FILE REFERENCE: 100086.409
; CURRENT APPLICATION NUMBER: US/09/185,908A
; CURRENT FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on mouse claudin-2 sequence
US-09-185-908-42

Query Match 60.9%; Score 14; DB 15; Length 8;
Best Local Similarity 25.0%; Pred. No. 5.1e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 1 WRTSSVVG 8

RESULT 2

US-09-185-908-163
; Sequence 163, Application US/09185908A
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C1
; CURRENT APPLICATION NUMBER: US/09/185,908A
; CURRENT FILING DATE: 1998-11-03
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 163
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on mouse claudin-2 sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide
US-09-185-908-163

Query Match 60.9%; Score 14; DB 15; Length 8;
Best Local Similarity 25.0%; Pred. No. 5.1e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 1 WRTSSVVG 8

RESULT 3

US-09-282-029-42
; Sequence 42, Application US/09282029
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C1
; CURRENT APPLICATION NUMBER: US/09/282,029
; CURRENT FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on mouse claudin-2 sequence
US-09-282-029-42

Query Match 60.9%; Score 14; DB 15; Length 8;
Best Local Similarity 25.0%; Pred. No. 5.1e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 1 WRTSSVVG 8

RESULT 4

US-09-282-029-163
; Sequence 163, Application US/09282029

; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C1
; CURRENT APPLICATION NUMBER: US/09/282,029
; CURRENT FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 163
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on mouse claudin-2 sequence
; FEATURE:
; OTHER INFORMATION: Cyclic Peptide
US-09-282-029-163

Query Match 60.9%; Score 14; DB 16; Length 8;
Best Local Similarity 25.0%; Pred. No. 5.1e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 1 WRTSSVVG 8

RESULT 5

US-09-282-029-394
; Sequence 394, Application US/09282029
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C1
; CURRENT APPLICATION NUMBER: US/09/282,029
; CURRENT FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 394
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative linear modulating agent based on
; OTHER INFORMATION: claudin-7 cell adhesion recognition sequence
US-09-282-029-394

Query Match 60.9%; Score 14; DB 16; Length 8;
Best Local Similarity 25.0%; Pred. No. 5.1e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
Db 1 WQMSVVG 8

RESULT 6

US-09-282-029-447
; Sequence 447, Application US/09282029
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C1
; CURRENT APPLICATION NUMBER: US/09/282,029
; CURRENT FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 479

RESULT 11

```
US-09-434-355-42
; SEQUENCE 42, Application US/09434355
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: James Mathew Symonds
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C2
; CURRENT APPLICATION NUMBER: US/09/434,355
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on mouse claudin-2 sequence
US-09-434-355-42

Query Match          60.9%; Score 14; DB 18; Length 8;
Best Local Similarity 25.0%; Pred. No. 5.1e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 1 WETSSYVG 8

RESULT 12
US-09-434-355-163
; SEQUENCE 163, Application US/09434355
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: James Mathew Symonds
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C2
; CURRENT APPLICATION NUMBER: US/09/434,355
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 163
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on mouse claudin-2 sequence
; OTHER INFORMATION: Cyclic Peptide
US-09-434-355-163

Query Match          60.9%; Score 14; DB 18; Length 8;
Best Local Similarity 25.0%; Pred. No. 5.1e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 1 WETSSYVG 8

RESULT 13
US-09-434-355-394
; SEQUENCE 394, Application US/09434355
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: James Mathew Symonds
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
```

```
; TITLE OF INVENTION: FUNCTIONS
; FILE REFERENCE: 100086.409C2
; CURRENT APPLICATION NUMBER: US/09/434,355
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 394
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative linear modulating agent based on
; OTHER INFORMATION: claudin-7 cell adhesion recognition sequence
US-09-434-355-394

Query Match          60.9%; Score 14; DB 18; Length 8;
Best Local Similarity 25.0%; Pred. No. 5.1e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 1 WOMSSYAG 8

RESULT 14
US-09-434-355-447
; SEQUENCE 447, Application US/09434355
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: James Mathew Symonds
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C2
; CURRENT APPLICATION NUMBER: US/09/434,355
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 447
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative linear modulating agent based on
; OTHER INFORMATION: claudin-7 cell adhesion recognition sequence
US-09-434-355-447

Query Match          60.9%; Score 14; DB 18; Length 8;
Best Local Similarity 25.0%; Pred. No. 5.1e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 1 WOMSSYAG 8

RESULT 15
US-09-434-355A-42
; SEQUENCE 42, Application US/09434355A
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: James Mathew Symonds
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C2
; CURRENT APPLICATION NUMBER: US/09/434,355A
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 8
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Product of
; OTHER INFORMATION: synthesis based on mouse claudin-2 sequence
US-09-434-355A-42

Query Match      60.9%; Score 14; DB 18; Length 8;
Best Local Similarity 25.0%; Pred. No. 5.1e+06;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
   |
Db 1 WRTSSYVG 8
```

Search completed: August 16, 2003, 14:55:10
Job time : 366 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:46:04 ; Search time 22 Seconds
(without alignments)
9.375 Million cell updates/sec

Title: US-09-185-908-1
Perfect score: 23
Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 121698 seqs, 25781089 residues

Total number of hits satisfying chosen parameters: 8075

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	11	47.8	4	1 PCT-US03-11768-56	Sequence 56, Appl
2	11	47.8	4	1 PCT-US02-30316-14	Sequence 14, Appl
3	11	47.8	4	1 PCT-US02-30316-15	Sequence 15, Appl
4	11	47.8	4	1 PCT-US02-30316-16	Sequence 16, Appl
5	11	47.8	4	1 PCT-US03-20997-4	Sequence 4, Appl
6	11	47.8	4	1 PCT-US03-21417-1	Sequence 1, Appl
7	11	47.8	4	5 US-09-341-590A-37	Sequence 37, Appl
8	11	47.8	4	5 US-09-341-590A-52	Sequence 52, Appl
9	11	47.8	4	5 US-09-882-781A-1	Sequence 1, Appl
10	11	47.8	4	5 US-09-882-781A-2	Sequence 2, Appl
11	11	47.8	4	5 US-09-882-781A-3	Sequence 3, Appl
12	11	47.8	4	5 US-09-882-781A-4	Sequence 4, Appl
13	11	47.8	4	5 US-09-882-781A-5	Sequence 5, Appl
14	11	47.8	4	5 US-09-882-781A-6	Sequence 6, Appl
15	11	47.8	4	5 US-09-882-781A-7	Sequence 7, Appl
16	11	47.8	4	5 US-09-882-781A-8	Sequence 8, Appl
17	11	47.8	4	5 US-09-882-781A-9	Sequence 9, Appl
18	11	47.8	4	5 US-09-882-781A-10	Sequence 10, Appl
19	11	47.8	4	5 US-09-882-781A-11	Sequence 11, Appl
20	11	47.8	4	5 US-09-882-781A-12	Sequence 12, Appl
21	11	47.8	4	5 US-09-882-781A-13	Sequence 13, Appl
22	11	47.8	4	5 US-09-882-781A-14	Sequence 14, Appl
23	11	47.8	4	5 US-09-882-781A-15	Sequence 15, Appl
24	11	47.8	4	5 US-09-882-781A-16	Sequence 16, Appl
25	11	47.8	4	5 US-09-882-781A-17	Sequence 17, Appl
26	11	47.8	4	5 US-09-882-781A-18	Sequence 18, Appl

Sequence 19, Appl
Sequence 20, Appl
Sequence 21, Appl
Sequence 22, Appl
Sequence 23, Appl
Sequence 31, Appl
Sequence 32, Appl
Sequence 25, Appl
Sequence 1, Appl
Sequence 74, Appl
Sequence 8, Appl
Sequence 9, Appl
Sequence 14, Appl
Sequence 15, Appl
Sequence 23, Appl
Sequence 29, Appl
Sequence 53, Appl
Sequence 409, Appl
Sequence 468, Appl

27 11 47.8 4 5 US-09-882-781A-19
28 11 47.8 4 5 US-09-882-781A-20
29 11 47.8 4 5 US-09-882-781A-21
30 11 47.8 4 5 US-09-882-781A-22
31 11 47.8 4 5 US-09-882-781A-23
32 11 47.8 4 5 US-09-882-781A-31
33 11 47.8 4 5 US-09-882-781A-32
34 11 47.8 4 5 US-09-714-712B-25
35 11 47.8 4 5 US-09-738-948B-1
36 11 47.8 4 6 US-10-464-071-74
37 11 47.8 4 6 US-10-450-740-8
38 11 47.8 4 6 US-10-450-740-9
39 11 47.8 4 6 US-10-463-699-14
40 11 47.8 4 6 US-10-463-699-15
41 11 47.8 4 6 US-10-380-565-23
42 11 47.8 4 6 US-10-281-092-29
43 11 47.8 4 6 US-10-281-092-53
44 11 47.8 4 6 US-10-121-024B-409
45 11 47.8 4 6 US-10-121-024B-468

ALIGNMENTS

RESULT 1

PCT-US03-11768-56

; Sequence 56, Application PC/TUS0311768

; GENERAL INFORMATION:

; APPLICANT: University of South Florida

; APPLICANT: Litman, Gary W.

; APPLICANT: Hawke, Noel A.

; APPLICANT: Yoder, Jeffrey A.

; APPLICANT: Eason, Donna D.

; TITLE OF INVENTION: B1VM (Basic, Immunoglobulin-Like Variable Motif-Containing) G

; FILE REFERENCE: USF-103X PCT

; CURRENT APPLICATION NUMBER: PCT/US03/11768

; NUMBER OF SEQ ID NOS: 64

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 56

; LENGTH: 4

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: M2 amino acid motif

; FEATURE:

; NAME/KEY: MISC_FEATURE

; LOCATION: (2)..(2)

; OTHER INFORMATION: Xaa = Tyr or Phe

; FEATURE:

; NAME/KEY: MISC_FEATURE

; LOCATION: (4)..(4)

; OTHER INFORMATION: Xaa = Gln or His

PCT-US03-11768-56

Query Match 47.8%; Score 11; DB 1; Length 4;

Best Local Similarity 100.0%; Pred. No. 1e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1

Db 1 W 1

RESULT 2

PCT-US02-30316-14

; Sequence 14, Application PC/TUS0230316

; GENERAL INFORMATION:

; APPLICANT: Pastan, Ira H.

; APPLICANT: Salvatore, Giuliana

; APPLICANT: Beers, Richard

; APPLICANT: Kreitman, Robert J.

```

; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by The Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mutated Anti-CD22 Antibodies With Increased Affinity to
; TITLE OF INVENTION: CD22-Expressing Leukemia Cells
; FILE REFERENCE: 015280-438100PC
; CURRENT APPLICATION NUMBER: PCT/US02/30316
; CURRENT FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: US 60/325,360
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mutated RFB4
; OTHER INFORMATION: V-H CDR3 CD22 target region
PCT-US02-30316-14

```

```

Query Match 47.8%; Score 11; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05; 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 W 1
   |
Db 4 W 4

```

```

RESULT 3
PCT-US02-30316-15
; Sequence 15, Application PC/TUS0230316
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira H.
; APPLICANT: Salvatore, Giuliana
; APPLICANT: Kreitman, Robert J.
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by The Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mutated Anti-CD22 Antibodies With Increased Affinity to
; TITLE OF INVENTION: CD22-Expressing Leukemia Cells
; FILE REFERENCE: 015280-438100PC
; CURRENT APPLICATION NUMBER: PCT/US02/30316
; CURRENT FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: US 60/325,360
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mutated RFB4
; OTHER INFORMATION: V-H CDR3 CD22 target region
PCT-US02-30316-15

```

```

Query Match 47.8%; Score 11; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05; 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 W 1
   |
Db 4 W 4

```

```

RESULT 4
PCT-US02-30316-16
; Sequence 16, Application PC/TUS0230316
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira H.

```

```

; APPLICANT: Salvatore, Giuliana
; APPLICANT: Beers, Richard
; APPLICANT: Kreitman, Robert J.
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by The Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mutated Anti-CD22 Antibodies With Increased Affinity to
; TITLE OF INVENTION: CD22-Expressing Leukemia Cells
; FILE REFERENCE: 015280-438100PC
; CURRENT APPLICATION NUMBER: PCT/US02/30316
; CURRENT FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: US 60/325,360
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:mutated RFB4
; OTHER INFORMATION: V-H CDR3 CD22 target region
PCT-US02-30316-16

```

```

Query Match 47.8%; Score 11; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05; 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 W 1
   |
Db 4 W 4

```

```

RESULT 5
PCT-US03-20997-4
; Sequence 4, Application PC/TUS0320997
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Wade, EDRIS
; TITLE OF INVENTION: CASPASE 9 ACTIVATION AND USES THEREFOR
; FILE REFERENCE: AM101006 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/20997
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 10/191,254
; PRIOR FILING DATE: 2002-07-08
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Substrate
PCT-US03-20997-4

```

```

Query Match 47.8%; Score 11; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05; 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 W 1
   |
Db 1 W 1

```

```

RESULT 6
PCT-US03-21417-1
; Sequence 1, Application PC/TUS0321417
; GENERAL INFORMATION:
; APPLICANT: Sharma, Shubh D.
; APPLICANT: Shadiack, Annette M.
; APPLICANT: Yang, Wei
; APPLICANT: Rajpurohit, Ramesh
; TITLE OF INVENTION: Peptide Compositions for Treatment of Sexual Dysfunction

```



```

Query Match      47.8%; Score 11; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 W 1
Db      2 W 2

RESULT 11
US-09-882-781A-3
; Sequence 3, Application US/09882781A
; GENERAL INFORMATION:
; APPLICANT: Kuhner, Carol H.
; TITLE OF INVENTION: Chemically-Modified Peptides, Compositions, And Methods Of
; FILE REFERENCE: HER0050
; CURRENT APPLICATION NUMBER: US/09/882,781A
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2000-06-16
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Novel Sequence
US-09-882-781A-3

Query Match      47.8%; Score 11; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 W 1
Db      2 W 2

RESULT 12
US-09-882-781A-4
; Sequence 4, Application US/09882781A
; GENERAL INFORMATION:
; APPLICANT: Kuhner, Carol H.
; TITLE OF INVENTION: Chemically-Modified Peptides, Compositions, And Methods Of
; FILE REFERENCE: HER0050
; CURRENT APPLICATION NUMBER: US/09/882,781A
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2000-06-16
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Novel Sequence
US-09-882-781A-4

Query Match      47.8%; Score 11; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 W 1
Db      2 W 2

RESULT 13
US-09-882-781A-5
; Sequence 5, Application US/09882781A
; GENERAL INFORMATION:
; APPLICANT: Kuhner, Carol H.
; TITLE OF INVENTION: Chemically-Modified Peptides, Compositions, And Methods Of
; FILE REFERENCE: HER0050
; CURRENT APPLICATION NUMBER: US/09/882,781A
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2000-06-16
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Novel Sequence
US-09-882-781A-5

Query Match      47.8%; Score 11; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 W 1
Db      2 W 2

RESULT 14
US-09-882-781A-6
; Sequence 6, Application US/09882781A
; GENERAL INFORMATION:
; APPLICANT: Kuhner, Carol H.
; TITLE OF INVENTION: Chemically-Modified Peptides, Compositions, And Methods Of
; FILE REFERENCE: HER0050
; CURRENT APPLICATION NUMBER: US/09/882,781A
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2000-06-16
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Novel Sequence
US-09-882-781A-6

Query Match      47.8%; Score 11; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 W 1
Db      4 W 4

RESULT 15

```

US-09-882-781A-7
; Sequence 7, Application US/09882781A
; GENERAL INFORMATION:
; APPLICANT: Kuhnert, Carol H.
; APPLICANT: Romesser, James A.
; TITLE OF INVENTION: Chemically-Modified Peptides, Compositions, And Methods Of
; TITLE OF INVENTION: Production And Use
; FILE REFERENCE: HER0050
; CURRENT APPLICATION NUMBER: US/09/882,781A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,441
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: PCT/US01/19400
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Novel Sequence
US-09-882-781A-7

Query Match 47.8%; Score 11; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
Db 4 W 4

Search completed: August 16, 2003, 14:55:39
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:44:04 ; Search time 28 seconds
(without alignments)
12.089 Million cell updates/sec

Title: US-09-185-908-1
Perfect score: 23
Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 61971

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	56.5	8	4 US-09-324-541-11	Sequence 11, Appl
2	12	52.2	8	4 US-08-469-260A-564	Sequence 564, App
3	12	52.2	8	4 US-08-488-446-564	Sequence 564, App
4	12	52.2	8	4 US-08-467-344A-564	Sequence 564, App
5	11	47.8	2	1 US-08-415-099A-1	Sequence 1, Appl
6	11	47.8	2	2 US-08-451-822A-19	Sequence 19, Appl
7	11	47.8	3	1 US-07-858-842-5	Sequence 5, Appl
8	11	47.8	3	1 US-07-960-636B-4	Sequence 4, Appl
9	11	47.8	3	1 US-08-165-545-8	Sequence 8, Appl
10	11	47.8	3	1 US-08-165-545-8	Sequence 8, Appl
11	11	47.8	3	1 US-08-165-545-12	Sequence 12, Appl
12	11	47.8	3	1 US-08-305-768-27	Sequence 27, Appl
13	11	47.8	3	1 US-08-256-771-13	Sequence 13, Appl
14	11	47.8	3	1 US-08-256-771-17	Sequence 17, Appl
15	11	47.8	3	1 US-08-256-771-21	Sequence 21, Appl
16	11	47.8	3	1 US-08-433-037-14	Sequence 14, Appl
17	11	47.8	3	1 US-08-440-504-1	Sequence 1, Appl
18	11	47.8	3	1 US-08-381-984-13	Sequence 13, Appl
19	11	47.8	3	1 US-08-381-984-17	Sequence 17, Appl
20	11	47.8	3	1 US-08-381-984-21	Sequence 21, Appl
21	11	47.8	3	2 US-08-871-163-27	Sequence 27, Appl
22	11	47.8	3	2 US-08-685-589A-89	Sequence 89, Appl
23	11	47.8	3	2 US-08-685-589A-90	Sequence 90, Appl
24	11	47.8	3	2 US-08-685-589A-92	Sequence 92, Appl
25	11	47.8	3	2 US-08-685-589A-93	Sequence 93, Appl
26	11	47.8	3	2 US-08-685-589A-94	Sequence 94, Appl
27	11	47.8	3	3 US-08-767-903-27	Sequence 27, Appl

28 11 47.8 3 3 US-08-590-897A-25 Sequence 25, Appl
29 11 47.8 3 3 US-08-433-613-1 Sequence 1, Appl
30 11 47.8 3 3 US-09-461-697-291 Sequence 291, App
31 11 47.8 3 3 US-09-461-697-405 Sequence 405, App
32 11 47.8 3 4 US-08-871-561-44 Sequence 44, Appl
33 11 47.8 3 4 US-09-167-513-6 Sequence 6, Appl
34 11 47.8 3 4 US-09-149-476-525 Sequence 525, App
35 11 47.8 3 4 US-09-321-932B-44 Sequence 44, Appl
36 11 47.8 3 5 PCT-US95-11724-27 Sequence 27, Appl
37 11 47.8 4 1 US-07-801-388-6 Sequence 6, Appl
38 11 47.8 4 1 US-08-061-065-6 Sequence 6, Appl
39 11 47.8 4 1 US-07-938-436-1 Sequence 1, Appl
40 11 47.8 4 1 US-07-932-200-1 Sequence 1, Appl
41 11 47.8 4 1 US-08-159-617-7 Sequence 7, Appl
42 11 47.8 4 1 US-07-841-997A-32 Sequence 32, Appl
43 11 47.8 4 1 US-08-165-545-3 Sequence 3, Appl
44 11 47.8 4 1 US-08-165-545-6 Sequence 6, Appl
45 11 47.8 4 1 US-08-165-545-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-324-541-11
; Sequence 11, Application US/09324541
; Patent No. 6391855
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: JUNCTIONAL ADHESION MOLECULE-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.412
; CURRENT APPLICATION NUMBER: US/09/324,541
; CURRENT FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fab fragment directed against claudin cell
; OTHER INFORMATION: adhesion recognition sequence
US-09-324-541-11

Query Match 56.5%; Score 13; DB 4; Length 8;
Best Local Similarity 25.0%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
Db 1 WKIVSYAG 8

RESULT 2
US-08-469-260A-564
; Sequence 564, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIUK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;; STREET: 100 ABBOTT PARK ROAD
;; CITY: ABBOTT PARK
;; STATE: IL
;; COUNTRY: USA
;; ZIP: 60064-3500
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/469,260A
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/424,550
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: POREMBSKI, PRISCILLA E.
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 5527.PC.01
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-938-2623
;; INFORMATION FOR SEQ ID NO: 564:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 8 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-469-260A-564

Query Match 52.2%; Score 12; DB 4; Length 8;
Best Local Similarity 25.0%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 1 WHSGLGG 8

RESULT 3
US-08-488-446-564
; Sequence 564, Application US/08488446
; Patent No. 6558898
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/488,446
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/424,550
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: POREMBSKI, PRISCILLA E.
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 5527.PC.01
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-938-2623
;; INFORMATION FOR SEQ ID NO: 564:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 8 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-488-446-564

Query Match 52.2%; Score 12; DB 4; Length 8;
Best Local Similarity 25.0%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 1 WHSGLGG 8

RESULT 4
US-08-467-344A-564
; Sequence 564, Application US/08467344A
; Patent No. 6586568
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,344A
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/424,550
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: POREMBSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 708-937-6365

TELEFAX: 708-938-2623

INFORMATION FOR SEQ ID NO: 564:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 564:

US-08-467-344A-564

Query Match 52.28; Score 12; DB 4; Length 8;
Best Local Similarity 25.08; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 6; Indels 0;

QY 1 WXXXXXXG 8

DD 1 WHSGLIGG 8

RESULT 5

US-08-415-099A-1

Sequence 1, Application US/08415099A

Patent No. 5789384

GENERAL INFORMATION:

APPLICANT: Kravinson, V.Kh., Sery, S.V. and Morozov, V.G.

TITLE OF INVENTION: "Pharmaceutical Dipeptide Compositions and Methods of Use Thereof"

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cytoven

STREET: 10230 N.E. Points Drive, Suite 530

CITY: Kirkland

STATE: Washington

COUNTRY: USA

ZIP: 98033

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette-3.25 inch, 1.44Mb storage

COMPUTER: IBM PC/486 Compatible

OPERATING SYSTEM: MS-DOS 5.01

SOFTWARE: Word for Windows 6.0-t

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/415,099A

FILING DATE: 31-MAR-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/271,386

FILING DATE: 06-JUL-1994

APPLICATION NUMBER: 08/026,341

FILING DATE: March 4, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Sundmo, John, S.

REGISTRATION NUMBER: 34,446

REFERENCE/DOCKET NUMBER: 15548-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 1-206-889-5804(direct)

TELEFAX: 1-206-822-3644

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2 amino acids

TYPE: AMINO

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

US-08-415-099A-1

Query Match 47.88; Score 11; DB 1; Length 2;
Best Local Similarity 100.08; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0;

QY 1 W 1

DD 2 W 2

RESULT 6

US-08-451-822A-19

Sequence 19, Application US/08451822A

Patent No. 5863888

GENERAL INFORMATION:

APPLICANT: Dionne, Craig A

APPLICANT: Crumley, Greg

APPLICANT: Jaye, Michael C

APPLICANT: Schlensing, Joseph

TITLE OF INVENTION: Fibroblast Growth Factor Receptors

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Legal Department

STREET: 500 Arcola Road

CITY: Collegeville

STATE: PA

COUNTRY: USA

ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/451,822A

FILING DATE: 26-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/323,430

FILING DATE: 14-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/934,372

FILING DATE: 21-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/549,587

FILING DATE: 06-JUL-1990

ATTORNEY/AGENT INFORMATION:

NAME: Savitzky, Martin

REGISTRATION NUMBER: 29,699

REFERENCE/DOCKET NUMBER: A0496E

TELECOMMUNICATION INFORMATION:

TELEPHONE: (610) 454-3816

TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 2 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-451-822A-19

Query Match 47.88; Score 11; DB 2; Length 2;
Best Local Similarity 100.08; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0;

QY 1 W 1

DD 1 W 1

RESULT 7

US-07-858-842-5

Sequence 5, Application US/07858842

Patent No. 5314807

GENERAL INFORMATION:

APPLICANT: Yoshikawa, Masaaki

APPLICANT: Yokoyama, Keiichi

APPLICANT: Hasegawa, Masayasu

APPLICANT: Yasumoto, Ryouichi

APPLICANT: Fujita, Hiroyuki

1 TITLE OF INVENTION: METHOD FOR PRODUCING AN ANGIOTENSIN
2 NUMBER OF SEQUENCES: 5
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: ARMSTRONG & KUBOVCIK
5 STREET: 1725 K Street N.W., Suite 1000
6 CITY: Washington
7 STATE: District of Columbia
8 COUNTRY: United States of America
9 ZIP: 20006
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Floppy disk
12 COMPUTER: IBM PC compatible
13 OPERATING SYSTEM: PC-DOS/MS-DOS
14 SOFTWARE: Patentin Release #1.0, Version #1.25
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/07/858,842
17 FILING DATE: 19920327
18 CLASSIFICATION: 435
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: JP 142283
21 FILING DATE: 29-MAR-1991
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: JP 298060
24 FILING DATE: 17-OCT-1991
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Gormley, Mary E.
27 REGISTRATION NUMBER: 34, 409
28 REFERENCE/DOCKET NUMBER: 920247
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (202) 659-2930
31 TELEFAX: (202) 887-0357
32 TELEX: 440142
33 INFORMATION FOR SEQ ID NO: 5:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 3 amino acids
36 TYPE: AMINO ACID
37 TOPOLOGY: linear
38 MOLECULE TYPE: peptide
39 HYPOTHEICAL: NO
40 US-07-858-842-5

Query Match 47.8% Score 11; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
Db 3 W 3

RESULT 8
US-07-960-636B-4
; Sequence 4, Application US/07960636B
; Patent No. 5369015
; GENERAL INFORMATION:
; APPLICANT: Yoshikawa, Masaaki
; APPLICANT: Yokoyama, Keiichi
; APPLICANT: Hasegawa, Masayasu
; APPLICANT: Yasumoto, Ryouichi
; APPLICANT: Fujita, Hiroyuki
; TITLE OF INVENTION: Method for Producing an Angiotensin
; TITLE OF INVENTION: Converting Enzyme Inhibitor-Containing Composition
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westernman, Hattori, McLeand &
; ADDRESSEE: Naughton
; STREET: 1725 K Street, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

1 COMPUTER: IBM PC compatible
2 OPERATING SYSTEM: PC-DOS/MS-DOS
3 SOFTWARE: Patentin Release #1.0, Version #1.25
4 CURRENT APPLICATION DATA:
5 APPLICATION NUMBER: US/07/960,636B
6 FILING DATE: 14-OCT-1992
7 CLASSIFICATION: 435
8 PRIOR APPLICATION DATA:
9 APPLICATION NUMBER: JP 298061/1991
10 FILING DATE: 17-OCT-1991
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Tockman, Albert
13 REGISTRATION NUMBER: 19722
14 REFERENCE/DOCKET NUMBER: P1161-2679-A920798
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: (202) 659-2930
17 TELEFAX: (202) 887-0357
18 INFORMATION FOR SEQ ID NO: 4:
19 SEQUENCE CHARACTERISTICS:
20 LENGTH: 3 amino acids
21 TYPE: amino acid
22 TOPOLOGY: linear
23 MOLECULE TYPE: peptide
24 ORIGINAL SOURCE:
25 ORGANISM: Mammalian meat, fish, crustaceans
26 US-07-960-636B-4

Query Match 47.8% Score 11; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 W 1
|
Db 3 W 3

RESULT 9
US-08-165-545-4
; Sequence 4, Application US/08165545
; Patent No. 5424396
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptide and
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Displaywrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/165,545
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/871,981
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 3 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-165-545-4

Query Match 47.8%; Score 11; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 W 1
Db 1 W 1

RESULT 10
US-08-165-545-8
Sequence 8, Application US/08165545
Patent No. 5424396
GENERAL INFORMATION:
APPLICANT: Mamoru TOMITA et al.
TITLE OF INVENTION: Antimicrobial Peptide and
TITLE OF INVENTION: Antimicrobial Agent
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS
SOFTWARE: DisplayWrite
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/165,545
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/871,981
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM:
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
ORGANELLE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
MAP POSITION:
UNITS:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-165-545-8

Query Match 47.8%; Score 11; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 W 1
Db 2 W 2

RESULT 11

US-08-165-545-12
; Sequence 12, Application US/08165545
; Patent No. 5424396
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: Antimicrobial Peptide and
; TITLE OF INVENTION: Antimicrobial Agent
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Displaywrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/165,545
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/871,981
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:

US-08-165-545-12
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-08-165-545-12
Query Match 47.8%; Score 11; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 W 1
Db 2 W 2
RESULT 12
US-08-305-768-27
; Sequence 27, Application US/08305768
; Patent No. 5602097
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Antifungal Peptides
; NUMBER OF SEQUENCES: 31
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/305,768
; FILING DATE: 12-SEPT-1994
; CLASSIFICATION: 514
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-305-768-27
Query Match 47.8%; Score 11; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 W 1
Db 1 W 1
RESULT 13
US-08-256-771-13
; Sequence 13, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; TITLE OF INVENTION: PRODUCTS THEREWITH
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/256,771
;; FILING DATE: July 22, 1994
;; CLASSIFICATION: 514
;; PRIOR APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren M. Cheek, Jr.
;; REGISTRATION NUMBER: 33,367
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-8850
;; TELEFAX:
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-256-771-13

Query Match 47.8%; Score 11; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 W 1
|
Db 1 W 1

RESULT 14
US-08-256-771-17
; Sequence 17, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3 amino acids
; TYPE: amino acid

;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-256-771-17

Query Match 47.8%; Score 11; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 W 1
|
Db 2 W 2

RESULT 15
US-08-256-771-21
; Sequence 21, Application US/08256771
; Patent No. 5656591
; GENERAL INFORMATION:
; APPLICANT: Mamoru TOMITA et al.
; TITLE OF INVENTION: ANTIMICROBIAL AGENTS AND METHOD FOR TREATING
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,771
; FILING DATE: July 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-256-771-21

Query Match 47.8%; Score 11; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 W 1
|
Db 2 W 2

Search completed: August 16, 2003, 14:48:58
Job time : 29 secs

```

! FINDPATTERNS on swp:* allowing 0 mismatches
!
1 1 W(K,R)XX(S,A)(Y,F)XG August 15, 2003 10:47 ..
ADH3_ECOLI ck: 7661 len: 369 ! P25437 escherichia coli. alcohol dehydrogen
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE
309: VTGRV WKSFAFGG VKGRS
ADH3_HAEIN ck: 2043 len: 378 ! P44557 haemophilus influenzae. putative alc
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE
318: VTGRV WKSFAFGG VKGRS
ADH3_PASPI ck: 798 len: 369 ! P39450 pasteurella piscicida (photobacteri
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE
309: VTGRV WKSFAFGG VKGRS
ADH3_GADMO ck: 1925 len: 375 ! P81600 gadus morhua (atlantic cod). alcohol
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE
315: VTGRV WKSFAFGG VKGRS
ADH1_RHOSH ck: 2905 len: 376 ! P72324 rhodobacter sphaeroides (rhodospseud
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG ARGRT
316: VTGRV WKSFAFGG VKGRS
ADH1_GADMO ck: 1722 len: 375 ! P81601 gadus morhua (atlantic cod). alcohol
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG YKSVE
315: VTGRV WKSFAFGG VKGRS
ADH1_ARATH ck: 5269 len: 379 ! Q96533 arabidopsis thaliana (mouse-ear cres
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG FKSRT
317: VTGRV WKSFAFGG VKGRS
ADH1_CAEEL ck: 9402 len: 384 ! Q17335 caenorhabditis elegans. alcohol dehyd
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE
321: VTGRV WKSFAFGG VKGRS
ADH1_DROME ck: 8245 len: 378 ! P46415 drosophila melanogaster (fruit fly)
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE
318: VVGRV WKSFAFGG WKSVE
ADH1_HORSE ck: 6584 len: 373 ! P19854 equus caballus (horse). alcohol dehyd
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE
313: VTGRV WKSFAFGG VKGRS
ADH1_ORYZA ck: 4843 len: 381 ! P93436 oryza sativa (rice). alcohol dehyd
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG FKSRS
319: VTGRV WKSFAFGG VKGRS
ADH1_PEA ck: 405 len: 378 ! P80572 pisum sativum (garden pea). alcohol
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG FKSRS
316: VTGRV WKSFAFGG VKGRS
ADH1_RABIT ck: 5825 len: 373 ! O19053 oryctolagus cuniculus (rabbit).
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE
313: VTGRV WKSFAFGG VKGRS
ADH1_RAT ck: 6609 len: 373 ! P12711 rattus norvegicus (rat). alcohol
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE
313: VTGRV WKSFAFGG VKGRS
ADH1_SPAD ck: 6805 len: 376 ! P79896 sparus aurata (gilthead sea brea
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE
316: VTGRV WKSFAFGG VKGRS
ADH1_UROHA ck: 5507 len: 373 ! P80467 uromastix hardwickii (indian spi

```

1	313: VTGFT	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGTAFGG	WKSVE		29: CALPM	WRVTAFIG	SSIIT
1	AP54_YEAST	ck: 8779 len: 475	! Q00776 saccharomyces cerevisiae (baker's yeast)		CLD3_HUMAN	ck: 1473 len: 220	! O15551 homo sapiens (human). claudin-3 (
1	382: KSAIL	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKIRSPFG	GKEYS		29: CALPM	WRVSAFAG	SSIIT
1	AQL1_THEAQ	ck: 2790 len: 513	! P08594 thermus aquaticus. aqualysin 1 precursor		CLD3_MOUSE	ck: 7312 len: 219	! Q920g9 mus musculus (mouse). claudin-3 (
1	495: AGYYL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRIYAYSG	SGMYE		29: CALPM	WRVSAFAG	SSIIT
1	ASSY_LEPIN	ck: 6380 len: 403	! Q8eyp7 leptospira interrogans. argininosuccinate lyase		CLD3_RAT	ck: 8372 len: 219	! Q63400 rattus norvegicus (rat). claudin-3 (
1	149: TIAP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG WRIWSFGG	RSDLI		29: CALPM	WRVSAFAG	SSIIT
1	CLD1_HUMAN	ck: 2066 len: 211	! O95832 homo sapiens (human). claudin-1 (sequence)		CLD4_CERAE	ck: 6200 len: 209	! O19005 cercopithecus aethiops (green monkey)
1	30: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRIYSYAG	DNIVT		30: CALPM	WRVTAFIG	SNIVT
1	CLD1_MOUSE	ck: 2898 len: 211	! O88551 mus musculus (mouse). claudin-1. 9/2		CLD4_HUMAN	ck: 6154 len: 209	! O14493 homo sapiens (human). claudin-4 (
1	30: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	DNIVT		30: CALPM	WRVTAFIG	SNIVT
1	CLD1_RAT	ck: 2787 len: 211	! P56745 rattus norvegicus (rat). claudin-1. 1		CLD4_MOUSE	ck: 1978 len: 210	! O35054 mus musculus (mouse). claudin-4 (
1	30: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	DNIVT		30: CALPM	WRVTAFIG	SNIVT
1	CLD2_CANFA	ck: 8125 len: 230	! Q95xm6 canis familiaris (dog). claudin-2. 2		CLD6_HUMAN	ck: 5150 len: 220	! P56747 homo sapiens (human). claudin-6 (
1	30: MLLPS	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTSSVVG	TSIVT		30: CALPM	WKVTAFIG	NSIVV
1	CLD2_HUMAN	ck: 8102 len: 230	! P57739 homo sapiens (human). claudin-2. 9/2		CLD6_MOUSE	ck: 3382 len: 219	! Q92262 mus musculus (mouse). claudin-6 (
1	30: MLLPS	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WRTSSVVG	ASIVT		30: CALPM	WKVTAFIG	NSIVV
1	CLD2_MOUSE	ck: 6271 len: 230	! O88552 mus musculus (mouse). claudin-2. 5/2		CLD9_HUMAN	ck: 8813 len: 217	! O95484 homo sapiens (human). claudin-9.
1	30: MLLPN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTSSVVG	ASIVT		30: CALPL	WKVTAFIG	NSIVV
1	CLD3_CANFA	ck: 3808 len: 218	! Q95xm5 canis familiaris (dog). claudin-3. 2		CLD9_MOUSE	ck: 9358 len: 217	! Q920s7 mus musculus (mouse). claudin-9.
1		W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG			30: CALPL	WKVTAFIG	NSIVV

1 LOXL_MOUSE ck: 3173 len: 662 ! P39654 mus musculus (mouse). arachidonate 1
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
222: RVRNS WKEDAFGG YOFLN

1 MGTA_THEMA ck: 982 len: 441 ! P80099 thermotoga maritima. 4-alpha-glucan
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
348: EGOTF WKWPAYNG PFSGI

1 MYBB_XENLA ck: 2650 len: 743 ! P52551 xenopus laevis (african clawed frog)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
705: PMTAA WKTVAEGG SODQM

1 OM25_BRUAB ck: 5845 len: 213 ! Q44664 brucella abortus. 25 kda outer-membr
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
74: IKPDD WKAGAFAG WNFQQ

1 OM25_BRUCA ck: 6016 len: 213 ! Q45110 brucella canis. 25 kda outer-membran
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
74: IKPDD WKAGAFAG WNFQK

1 OM25_BRUME ck: 5809 len: 213 ! Q45321 brucella melitensis. 25 kda outer-me
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
74: IKPDD WKAGAFAG WNFQQ

1 OM25_BRUNE ck: 5851 len: 213 ! Q45326 brucella neotomae. 25 kda outer-memb
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
74: IKPDD WKAGAFAG WNFQQ

1 OM25_BRUOV ck: 2394 len: 201 ! Q45335 brucella ovis. 25 kda outer-membrane
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
74: IKPDD WKAGAFAG WNFQQ

1 OM25_BRUSU ck: 5911 len: 213 ! Q45689 brucella suis. 25 kda outer-membrane
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
74: IKPDD WKAGAFAG WNFQQ

1 R24E_METJA ck: 6435 len: 70 ! P54064 methanococcus jannaschii. 50s ribos
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
4: MPE WRTCSFCG YEIEP

1 RHG7_HUMAN ck: 9918 len: 1,091 ! Q96qb1 homo sapiens (human). rho-gtpase-act

1 W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
401: NGSVN WRTGSFHG PGHIS

1 RHG7_MOUSE ck: 9678 len: 1,092 ! Q9r0z9 mus musculus (mouse). rho-gtpase-
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
402: NGSVN WRTGSFHG PGHLS

1 RHG7_RAT ck: 1766 len: 1,091 ! Q63744 r rho-gtpase-activating protein 7
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
401: NGSVN WRTGSFHG PGHLS

1 SERA_MYCLE ck: 421 len: 528 ! O33116 mycobacterium leprae. d-3-phospho
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
129: LRAHI WKSSFSFG TEIFG

1 SERA_MYCTU ck: 1055 len: 528 ! O53243 mycobacterium tuberculosis. d-3-p
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
129: LREHT WKSSFSFG TEIFG

1 SPS2_MOUSE ck: 9022 len: 452 ! P97364 mus musculus (mouse). selenide,wa
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
52: GFSPS WRUTSFSFG MKGCG

1 SSAM_SALTY ck: 5313 len: 122 ! P74855 salmonella typhimurium. secretion
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
70: KALAN WKPAAFQG IPQL

1 STAD_LINUS ck: 845 len: 396 ! P32062 linum usitatissimum (flax) (linse
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
340: FLVGR WKVDAFTG LSGEG

1 UFD2_SCHFO ck: 3117 len: 1,010 ! Q9he05 schizosaccharomyces pombe (fissio
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
706: VERTI WKQAFYFG KLEQE

1 VA28_VARY ck: 405 len: 146 ! P33847 variola virus. protein a28. 10/20
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(Y)XG
73: DVKQK WRCVAYPG NGFVS

1 VAT_CAMVC ck: 3857 len: 159 ! P03549 cauliflower mosaic virus (strain
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
61: SLLGI WKINSYFG LSKDP

309: VTGRV	WRGSAFEG	VKGRS	
Q8X0U5	ck: 9212 len: 380	! Q8X0U5 neurospora crassa. probable alcohol	
	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG		
319: VTGRV	WRGSAFEG	VKGRS	
O74636	ck: 8858 len: 777	! O74636 fusarium oxysporum. transposase-like	
	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG		
402: ERFEI	WRKQSFIG	KLHNF	
Q96V39	ck: 6889 len: 380	! Q96V39 pichia angusta (yeast) (hansenula po	
	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG		
318: VTGRV	WRGCAFGG	IKGRT	
Q8J2V2	ck: 8797 len: 230	! Q8J2V2 gibberella zeae (fusarium graminearu	
	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG		
170: IDTDD	WRYSSFGG	AEPSL	
Q8J0F4	ck: 9502 len: 1,173	! Q8J0F4 penicillium citrinum. hmg-coa reduct	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG		
97: GPQNG	WKWQSFDD	DADVL	
Q8J0F1	ck: 4488 len: 380	! Q8J0F1 candida boidinii (yeast). formaldehy	
	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG		
318: VTGRV	WRGCAFGG	VKGRS	
Q8WZB3	ck: 404 len: 26,926	! Q8WZB3 homo sapiens (human). n2b-titin iso	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG		
18,126: NVTLK	WKKPAYDG	GSKIT	
Q8NFW8	ck: 3639 len: 434	! Q8NFW8 homo sapiens (human). cytidine mono	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG		
365: EMGLC	WKEVAYLG	NEVSD	
Q8WZ42	ck: 1298 len: 34,350	! Q8WZ42 homo sapiens (human). titin. 3/2003	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG		
25,550: NVTLK	WKKPAYDG	GSKIT	
Q8NEI1	ck: 587 len: 496	! Q8NEI1 homo sapiens (human). hypothetical p	
	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG		
4: MAL	WRGSAYAG	FLALA	

Q8TC92	ck: 6092 len: 643	! Q8TC92 homo sapiens (human). hypotheticala	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG		
632: TLEKR	WKLCAFEF	IKTT	
Q9NWE0	ck: 6755 len: 643	! Q9NWE0 homo sapiens (human). hypotheticala	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG		
632: TLEKR	WKLCAFEF	IKTT	
Q9NQZ0	ck: 2647 len: 434	! Q9NQZ0 homo sapiens (human). cmp-n-acety	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG		
365: EMGLC	WKEVAYLG	NEVSD	
Q9BRE6	ck: 3596 len: 497	! Q9BRE6 homo sapiens (human). similar to	
	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG		
4: MAL	WRGSAYAG	FLALA	
Q10466	ck: 227 len: 26,926	! Q10466 homo sapiens (human). titin, hea	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG		
18,126: NVTLK	WKKPAYDG	GSKIT	
Q96DQ2	ck: 9961 len: 811	! Q96DQ2 homo sapiens (human). hypotheticala	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG		
384: QINEN	WKRHSYAG	EQPET	
Q8N0N4	ck: 7283 len: 99	! Q8N0N4 branchiostoma floridae (florida 1	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG		
39: VTGRV	WKGTAFFG	WKSVE	
Q9NJD0	ck: 9928 len: 377	! Q9NJD0 branchiostoma floridae (florida 1	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG		
317: VTGRV	WKGTAFFG	WKSVE	
Q8WS90	ck: 6247 len: 377	! Q8WS90 ciona intestinalis. alcohol dehyd	
	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG		
317: VTGRV	WKGTAFFG	YKSVE	
Q9NE65	ck: 4708 len: 1,778	! Q9NE65 leishmania major. hypothetical 18	
	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG		
1,309: VRVAS	WRWSFFPG	QLHSL	
Q9NJC3	ck: 745 len: 377	! Q9NJC3 branchiostoma lanceolatum (common	
	W(K,R)XX(S,A)(Y,F)XG		

1	317: VTGRV	W(K)XX(A)(F)XG WKGTAFGG	WKSVD	Q9BJ33	ck: 1200 len: 377	! Q9bj33 branchiostoma floridae (florida lanc	1	Q9N356	ck: 7111 len: 317	! Q9n356 caenorhabditis elegans. hypothet
		W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKSVE					241: VVKMT	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	VLAFF
	317: VTGRV	WKGTAFGG							WKQSYDG	
1	Q9NFP2	ck: 1806 len: 432	! Q9nfp2 plasmodium falciparum. nima-related					Q9VA05	ck: 5184 len: 1,431	! Q9va05 drosophila melanogaster (fruit f
		W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	LKERE					318: VVIAT	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	GIDNS
	43: QEFPC	WKAISYRG							WKNVSEFAG	
1	Q965R0	ck: 8060 len: 554	! Q965r0 caenorhabditis elegans. hypothetical					Q8I5D5	ck: 9049 len: 1,057	! Q8i5d5 plasmodium falciparum (isolate :
		W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKSVE						W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	
	494: VTGRT	WKGTAFGG						43: QEFPC	WKAISYRG	LKERE
1	Q9BJ34	ck: 835 len: 377	! Q9bj34 branchiostoma floridae (florida lanc					Q95K71	ck: 1821 len: 414	! Q95k71 macaca fascicularis (crab eatin
		W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKSVE						W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	
	317: VTGRV	WKGTAFGG						403: TLEKR	WKLCAFEQ	IKTT
1	Q8NS89	ck: 6358 len: 377	! Q8ws89 ciona intestinalis. alcohol dehydro					Q28733	ck: 9453 len: 6,875	! Q28733 oryctolagus cuniculus (rabbit).
		W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	YKSVE						W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	
	317: VTGRT	WKGTAFGG						6,555: NVTLK	WKKPAYDG	GSKIT
1	Q18005	ck: 8316 len: 545	! Q18005 caenorhabditis elegans. hypothetical					Q9B394	ck: 8082 len: 255	! Q9b394 blackburnia palmae. cytochrome i
		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG	FSNGV						W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG	
	404: GLVGS	WRNSSEFSG						85: GYVLP	WQMSFWG	ATVIT
1	Q97363	ck: 7232 len: 313	! Q97363 bombyx mori (silk moth). lipopolysac					Q9GI47	ck: 9519 len: 348	! Q9gi47 daphniphyllum sp. qiu 94162. mai
		W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	PHDWN						W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	
	230: MIGSF	WKDVAFIC						77: ISAND	WQNSFLG	HNKNL
1	Q961U1	ck: 4884 len: 1,721	! Q961u1 drosophila melanogaster (fruit fly).					Q38144	ck: 957 len: 159	! Q38144 bacteriophage spp1. required fo
		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	GLKVS						W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG	
	253: SCMLN	WRPPSYDG						77: GVSQS	WETGAFES	QDGKR
1	Q9NEV9	ck: 7429 len: 327	! Q9nev9 manduca sexta (tobacco hawkmoth) (to					Q9B0E2	ck: 7287 len: 412	! Q9b0e2 staphylococcus aureus temperate
		W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	PHDWG						W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	
	232: MYGNF	WKDMAFVG						35: YDFSP	WKNRSFWG	VINNT
1	Q9VA47	ck: 9013 len: 8,971	! Q9va47 drosophila melanogaster (fruit fly).					Q8S9Y0	ck: 4807 len: 929	! Q8s9y0 oryza sativa (japonica cultivar
		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	GLKVS						W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG	
	7,429: SCMLN	WRPPSYDG						266: KEDCP	WRVHAYKG	KWNDY
1								Q8RUM3	ck: 5321 len: 542	! Q8rum3 oryza sativa (japonica cultivar

1	247: FGRCPC	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTLSYQG	KLEWV	Q9FW81	ck: 4017	len: 1,626	! Q9fw81 oryza sativa (rice). mutator-like tr	1	Q9LDE7	ck: 51	len: 938	! Q9lde7 oryza sativa (rice). est c28952(c		
1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q9FW81	ck: 4017	len: 1,626	! Q9fw81 oryza sativa (rice). mutator-like tr	1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY			
1	Q8LRF53	ck: 1421	len: 1,605	! Q8lr53 oryza sativa (japonica cultivar-grou	1	Q8LRF53	ck: 1421	len: 1,605	! Q8lr53 oryza sativa (japonica cultivar-grou	1	Q8W3H2	ck: 8786	len: 1,638	! Q8w3h2 oryza sativa (rice). mutator-like
1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q8LRF53	ck: 1421	len: 1,605	! Q8lr53 oryza sativa (japonica cultivar-grou	1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY			
1	Q8LQ12	ck: 4995	len: 1,592	! Q8lq12 oryza sativa (japonica cultivar-grou	1	Q8LQ12	ck: 4995	len: 1,592	! Q8lq12 oryza sativa (japonica cultivar-grou	1	Q9FGN6	ck: 275	len: 895	! Q9fgn6 arabidopsis thaliana (mouse-ear c
1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q8LQ12	ck: 4995	len: 1,592	! Q8lq12 oryza sativa (japonica cultivar-grou	1	621: RSQGQ	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKMVSFAG	LPHFT			
1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q9FRA2	ck: 2655	len: 1,011	! Q9fra2 oryza sativa (rice). similar to oryza	1	Q94LE7	ck: 4734	len: 883	! Q94le7 oryza sativa (rice). putative tra		
1	Q9FRA2	ck: 2655	len: 1,011	! Q9fra2 oryza sativa (rice). similar to oryza	1	Q9FRA2	ck: 2655	len: 1,011	! Q9fra2 oryza sativa (rice). similar to oryza	1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	VEKYH	
1	277: KDGCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWKDY	Q947Y7	ck: 3231	len: 2,421	! Q947y7 oryza sativa (rice). putative mutato	1	Q8SB37	ck: 39	len: 1,656	! Q8sb37 oryza sativa (rice). putative tra		
1	172: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q947Y7	ck: 3231	len: 2,421	! Q947y7 oryza sativa (rice). putative mutato	1	266: KEECP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY			
1	Q94B87	ck: 3033	len: 1,626	! Q94b87 oryza sativa (rice). putative mutato	1	Q94B87	ck: 3033	len: 1,626	! Q94b87 oryza sativa (rice). putative mutato	1	645: PQQIN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRPPSYFG	RDDLE	
1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q9AYG3	ck: 3386	len: 1,641	! Q9ayg3 oryza sativa (rice). mutator-like tr	1	Q8S5Y6	ck: 7546	len: 1,557	! Q8s5y6 oryza sativa (japonica cultivar-g		
1	Q9AYG3	ck: 3386	len: 1,641	! Q9ayg3 oryza sativa (rice). mutator-like tr	1	Q9AYG3	ck: 3386	len: 1,641	! Q9ayg3 oryza sativa (rice). mutator-like tr	1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	
1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q9LDW9	ck: 2865	len: 1,591	! Q9ldw9 oryza sativa (rice). est c28952(c629	1	Q8S696	ck: 8699	len: 1,530	! Q8s696 oryza sativa (japonica cultivar-g		
1	Q9LDW9	ck: 2865	len: 1,591	! Q9ldw9 oryza sativa (rice). est c28952(c629	1	Q9LDW9	ck: 2865	len: 1,591	! Q9ldw9 oryza sativa (rice). est c28952(c629	1	271: KDGCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWKDY	
1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q8W5M7	ck: 4750	len: 1,597	! Q8w5m7 oryza sativa (rice). putative mutato	1	Q8S1F8	ck: 8647	len: 1,080	! Q8s1f8 oryza sativa (japonica cultivar-g		
1	Q8W5M7	ck: 4750	len: 1,597	! Q8w5m7 oryza sativa (rice). putative mutato	1	Q8W5M7	ck: 4750	len: 1,597	! Q8w5m7 oryza sativa (rice). putative mutato	1	742: AVFDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRMFAFAG	AGDEQ	
1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q94FE8	ck: 695	len: 2,910	! Q94fb8 schizochytrium sp. atcc_20888. polyu	1	Q04892	ck: 2338	len: 530	! Q04892 nicotiana tabacum (common tobacco		
1	Q94FE8	ck: 695	len: 2,910	! Q94fb8 schizochytrium sp. atcc_20888. polyu	1	Q94FE8	ck: 695	len: 2,910	! Q94fb8 schizochytrium sp. atcc_20888. polyu	1	254: RGCSF	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRPDAFVG	GNDYY	
1	181: ANFPE	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG WRLDSFGG	FLGNV	Q94FE8	ck: 695	len: 2,910	! Q94fb8 schizochytrium sp. atcc_20888. polyu	1	Q64761	ck: 1756	len: 303	! Q64761 arabidopsis thaliana (mouse-ear c		

1	15: IPGPK	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRKVAYGG	MQICY	Q9LDA3	ck: 5721	len: 1,281	! Q9lda3 oryza sativa (rice). est c28952(c629	1	266: KEDCP	WRVHAYKG	KWNDY	Q8RYT2	ck: 5760	len: 1,110	! Q8ryt2 oryza sativa (japonica cultivar-
1	237: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q8LSZ5	ck: 7458	len: 602	! Q8lsz5 physcomitrella patens (moss). ferred	1	209: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q9SBA2	ck: 9787	len: 396	! Q9sba2 linum usitatissimum (flax) (lins
1	394: LIDPS	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKRSSFPG	VNPQK	Q8W062	ck: 1142	len: 1,604	! Q8w062 oryza sativa (rice), and oryza sativa	1	340: FLVGR	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKVDAFTG	LSGEG	Q94JG7	ck: 2586	len: 429	! Q94jg7 oryza sativa (rice), and oryza s
1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	O82014	ck: 9934	len: 396	! O82014 linum usitatissimum (flax) (linseed)	1	347: RDAMV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKVASFPG	GKDYM	O22715	ck: 2687	len: 428	! O22715 arabidopsis thaliana (mouse-ear
1	340: FLVGR	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKVDAFTG	LSGEG	Q8S521	ck: 6884	len: 198	! Q8s521 zea mays (maize). d-type cyclin (fra	1	346: KDALV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKIKSFPG	NKEYM	Q9SW79	ck: 8565	len: 274	! Q9sw79 triticum aestivum (wheat). alter
1	97: AIDWI	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKVHAYYG	FGPLT	Q8S211	ck: 4528	len: 1,353	! Q8s211 oryza sativa (japonica cultivar-grou	1	21: GRRRR	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRISYWG	IEQSK	Q8H504	ck: 6567	len: 1,597	! Q8h504 oryza sativa (japonica cultivar-
1	265: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q94E86	ck: 1244	len: 603	! Q94e86 oryza sativa (rice). b1045d11.16 pro	1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q8H8E2	ck: 1340	len: 779	! Q8h8e2 oryza sativa (japonica cultivar-
1	410: QTNLS	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKYAFQGG	KIYAA	Q9XE23	ck: 820	len: 955	! Q9xe23 oryza sativa (rice). est c28952(c629	1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q8H7V5	ck: 7087	len: 1,596	! Q8h7v5 oryza sativa (japonica cultivar
1	115: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q8LN97	ck: 3122	len: 1,536	! Q8ln97 oryza sativa (japonica cultivar-grou	1	277: KDGGP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWKDY	Q8H5S9	ck: 5613	len: 1,179	! Q8h5s9 oryza sativa (japonica cultivar
1	266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	Q8LN69	ck: 2184	len: 655	! Q8ln69 oryza sativa (japonica cultivar-grou	1	273: KDGGP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWKDY	Q8H525	ck: 8356	len: 1,727	! Q8h525 oryza sativa (japonica cultivar
1		W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG							266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY				

1	Q8H4Y6	ck: 2309	len: 268	! Q8h4y6 oryza sativa (japonica cultivar-group)	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG	ARRSV	243: FFRSF	WRAVSFSG	314: VTGRT	W(K)XX(A)(F)XG WKGTAFGG	WKSVE	
1	Q8H2L9	ck: 9400	len: 1,662	! Q8h2l9 oryza sativa (japonica cultivar-group)	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	KWNDY	266: KEDCP	WRVHAYKG	Q8C330	ck: 6000	len: 200	! Q8c330 mus musculus (mouse). cytidine mo
1	Q8H2L7	ck: 8873	len: 1,753	! Q8h2l7 oryza sativa (japonica cultivar-group)	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	KWNDY	266: KEDCP	WRVHAYKG	131: EMGLC	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKEVAYLG	NEVSD	
1	Q8CX38	ck: 3045	len: 265	! Q8gx38 arabidopsis thaliana (mouse-ear cres	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	RDDLE	210: PQQIN	WRPPSVFG	Q8BY11	ck: 241	len: 305	! Q8by11 mus musculus (mouse). putative hi
1	Q8GRQ7	ck: 3759	len: 903	! Q8grq7 oryza sativa (japonica cultivar-group)	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	KWNDY	266: KEECP	WRVHAYKG	198: QKQPA	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVSAFIG	KAQIS	
1	O88719	ck: 9067	len: 432	! O887l9 mus musculus (mouse). cmp-n-acetyline	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	NEVSD	363: EMGLC	WKEVAYLG	30: TLLPQ	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKRHSYAG	SNIII	
1	Q8R5B1	ck: 3258	len: 490	! Q8r5b1 mus musculus (mouse). similar to hyf	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	KAQIS	198: QKQPA	WKVGAYKG	Q8BXA6	ck: 890	len: 224	! Q8bxa6 mus musculus (mouse). similar to
1	Q8K2G7	ck: 8878	len: 432	! Q8k2g7 mus musculus (mouse). cytidine monoph	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	NEVSD	363: EMGLC	WKEVAYLG	Q8BWY1	ck: 675	len: 689	! Q8bwy1 mus musculus (mouse). hypothetica
1	Q99KK2	ck: 2584	len: 166	! Q99kk2 mus musculus (mouse). similar to cyt	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	NEVSD	97: EMGLC	WKEVAYLG	Q8BJ63	ck: 3056	len: 490	! Q8bj63 mus musculus (mouse). putative hi
1	Q9JHG0	ck: 2510	len: 197	! Q9jhg0 mus musculus (mouse). cbln3. 6/2001	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	FLIFP	184: NLLGG	WKYSFSG	Q8BJ58	ck: 6723	len: 398	! Q8bj58 mus musculus (mouse). putative hi
1	Q8C662	ck: 221	len: 374	! Q8c662 mus musculus (mouse). alcohol dehydra	W(K,R)XX(S,A)(Y,F)XG				198: QKQPA	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKVGAYKG	KAQIS	
1									Q8BJ31	ck: 3487	len: 307	! Q8bj31 mus musculus (mouse). putative hi
1									Q8BI12	ck: 3107	len: 490	! Q8bi12 mus musculus (mouse). putative hi
1									Q8BHR2	ck: 4917	len: 643	! Q8bhr2 mus musculus (mouse). hypothetica
1									632: TLEKR	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKLCAFEF	IKTT	
1									Q91TT0	ck: 2878	len: 138	! Q91tt0 tupai herpesvirus. t22.9. 10/200
1									67: TSVFV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG WRLCSFGG	GFVLG	

1	Q8V715	ck: 8144	len: 547	! Q8v715 swine calicivirus. capsid protein. 6	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	73: DVKQK WRCVAYFG	NGFVS
1	410: QDFNQ	WRLPAYGG	ALFNN					
1	Q8QNH4	ck: 1050	len: 284	! Q8qnh4 ectocarpus siliculosus virus. esv-1	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKFRATLG	Q90XR4	ck: 7881 len: 210 ! Q90xr4 brachydanio rerio (zebrafish) (d)
1	75: NCLKN	WKFRATLG	SGAHG			W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKVTAFIG	27: CALPM TNIVV	
1	Q9QQN6	ck: 7259	len: 479	! Q9qqn6 sugarcane yellow leaf virus. putative	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKVTAFIG	Q90XR8	ck: 4395 len: 215 ! Q90xr8 brachydanio rerio (zebrafish) (d)
1	129: ANLAG	WRAYATSG	CTISN			W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVSAFVG	30: CALPM ANIVT	
1	Q66159	ck: 2936	len: 159	! Q66159 cauliflower mosaic virus. orf ii. 6	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	Q90XR0	ck: 4854 len: 214 ! Q90xr0 brachydanio rerio (zebrafish) (d)
1	61: SLLGI	WKINSYFG	LSKDP			W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKVTAFIG	29: CALPM ANIVT	
1	Q83166	ck: 3141	len: 159	! Q83166 cauliflower mosaic virus. aphid trans	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	Q90XR9	ck: 2537 len: 209 ! Q90xr9 brachydanio rerio (zebrafish) (d)
1	61: SLLGI	WKINSYFG	LSKDP			W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKVTAFIG	30: CALPM ANIVT	
1	Q83179	ck: 3686	len: 159	! Q83179 cauliflower mosaic virus. orf ii. pro	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	Q9DE12	ck: 435 len: 214 ! Q9del2 xenopus laevis (african clawed f
1	61: SLLGI	WKINSYFG	LSKDP			W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKVTAFIG	30: CALPM NNIVV	
1	Q9JH75	ck: 3583	len: 479	! Q9jh75 sugarcane yellow leaf virus. putative	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRAYATSG	Q90XR6	ck: 1235 len: 211 ! Q90xr6 brachydanio rerio (zebrafish) (c
1	129: ANLAG	WRAYATSG	CTISN			W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKTSAFIG	30: IAIQO QNIIT	
1	Q9WI33	ck: 2957	len: 159	! Q9wi33 cauliflower mosaic virus. aphid trans	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	Q90XR2	ck: 5577 len: 218 ! Q90xr2 brachydanio rerio (zebrafish) (c
1	61: SLLGI	WKINSYFG	LSKDP			W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKVTAFIG	30: CALPM TNIVT	
1	Q83157	ck: 8743	len: 64	! Q83157 cauliflower mosaic virus. aphid acqu	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGTAFGG	Q90XD4	ck: 5678 len: 376 ! Q90xd4 brachydanio rerio (zebrafish) (c
1	7: SLLGI	WKINSYFG	LSKDP			W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKGTAFGG	316: VTGRT WKSVE	
1	Q83162	ck: 8680	len: 99	! Q83162 cauliflower mosaic virus. hypothetical	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	Q90XQ9	ck: 100 len: 214 ! Q90xq9 xenopus laevis (african clawed f
1	61: SLLGI	WKINSYFG	LSKDP			W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKVTAFIG	30: CAMPM NNIVV	
1	Q8V2M7	ck: 663	len: 146	! Q8v2m7 camelpox virus (strain cp-1). hypoth	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKVTAFIG	Q98SR2	ck: 3404 len: 214 ! Q98sr2 gallus gallus (chicken). claudir
							29: CALPM NNIVT	

1	Q90XQ8	ck: 8089	len: 210	! Q90xq8 brachydanio rerio (zebrafish) (danio)	1	VTGRV	ck: 8784	len: 538	! Q91le5 pseudomonas aeruginosa. probable
	W(K,R)XX(S,A)(Y,F)XG					W(K)XX(A)(F)XG			
	W(K)XX(S)(Y)XG					W(K)XX(A)(Y)XG			
30: TAMAE	WKMSYAG					WKIATDG			
	DNIIT					FRAQL			
1	Q80VX9	ck: 8315	len: 209	! Q80vx9 torpedo marmorata (marbled electric)	1	VRSER			
	W(K,R)XX(S,A)(Y,F)XG					W(K)XX(A)(Y)XG			
	W(K)XX(S)(F)XG					W(K)XX(A)(Y)XG			
90: QSSCA	WKOCSFG					W(K)XX(A)(Y)XG			
	VIQPH					Q9HY01	ck: 6470	len: 370	! Q9hy01 pseudomonas aeruginosa. alcohol d
1	Q90WG6	ck: 4720	len: 432	! Q90wg6 oncorhynchus mykiss (rainbow trout)	1	VTGRV			
	W(K,R)XX(S,A)(Y,F)XG					W(R)XX(A)(F)XG			
	W(K)XX(A)(Y)XG					VRGRS			
355: DKDLD	WKEVAYMG					Q9HT72	ck: 5958	len: 262	! Q9ht72 pseudomonas aeruginosa. permease
	NDAPD					W(K,R)XX(S,A)(Y,F)XG			
	Q8JIE6	ck: 8389	len: 1,019	! Q8jie6 gallus gallus (chicken). hira. 3/200	1	GSFVV			
1	W(K,R)XX(S,A)(Y,F)XG					W(R)XX(A)(Y)XG			
	W(K)XX(A)(Y)XG					DTLSH			
97: KLIWV	WKRAAYIG					Q9CP25	ck: 9951	len: 261	! Q9cp25 pasteurella multocida. hypothetic
	PSTVF					W(K,R)XX(S,A)(Y,F)XG			
	Q8AVG4	ck: 4992	len: 211	! Q8avg4 xenopus laevis (african clawed frog)	1	GAFVV			
1	W(K,R)XX(S,A)(Y,F)XG					W(R)XX(A)(Y)XG			
	W(K)XX(S)(F)XG					DTLAH			
30: IAIPO	WKSSFAG					Q9A956	ck: 9243	len: 613	! Q9a956 caulobacter crescentus. tonb-depe
	DAIIT					W(K,R)XX(S,A)(Y,F)XG			
1	Q8Q7G3	ck: 8999	len: 861	! Q8q7g3 human immunodeficiency virus 1. enve	1	KGDIW			
	W(K,R)XX(S,A)(Y,F)XG					W(R)XX(A)(Y)XG			
	W(R)XX(S)(Y)XG					FRPPT			
777: LASGI	WRVNSILG					Q9A5D4	ck: 8539	len: 369	! Q9a5d4 caulobacter crescentus. alcohol d
	LGLGI					W(K,R)XX(S,A)(Y,F)XG			
1	O67832	ck: 5280	len: 392	! O67832 aquifex aeolicus. hypothetical prote	1	VTGRV			
	W(K,R)XX(S,A)(Y,F)XG					W(K)XX(A)(F)XG			
	W(K)XX(S)(F)XG					ARGRT			
127: YVSYN	WKRSEFG					Q988W5	ck: 8644	len: 700	! Q988w5 rhizobium loti (mesorhizobium lot
	KDFED					W(K,R)XX(S,A)(Y,F)XG			
1	Q9KQB7	ck: 3098	len: 260	! Q9kqb7 vibrio cholerae. zinc abc transporte	1	ELLGP			
	W(K,R)XX(S,A)(Y,F)XG					W(R)XX(A)(F)XG			
	W(R)XX(A)(Y)XG					AGAGT			
28: GSFVV	WRMAYFG					Q987D8	ck: 5716	len: 412	! Q987d8 rhizobium loti (mesorhizobium lot
	DTLAH					W(K,R)XX(S,A)(Y,F)XG			
	Q9KPP6	ck: 9551	len: 1,208	! Q9kpp6 vibrio cholerae. exodeoxyribonuclea	1	NIERA			
1	W(K,R)XX(S,A)(Y,F)XG					W(R)XX(S)(F)XG			
	W(R)XX(S)(Y)XG					QTDAA			
906: AIDRR	WRVTSYG					Q984R5	ck: 9706	len: 219	! Q984r5 rhizobium loti (mesorhizobium lot
	LVMQS					W(K,R)XX(S,A)(Y,F)XG			
1	Q9KCG9	ck: 2788	len: 540	! Q9kcg9 bacillus halodurans. d-3-phosphoglyc	1	W(R)XX(S)(Y)XG			
	W(K,R)XX(S,A)(Y,F)XG					W(R)XX(S)(Y)XG			
	W(K)XX(A)(F)XG					KQORI			
141: IKAGE	WKRKAFQG					Q983F3	ck: 2074	len: 344	! Q983f3 rhizobium loti (mesorhizobium lot
	TELFG					W(K,R)XX(S,A)(Y,F)XG			
1	Q9JRB0	ck: 788	len: 378	! Q9jrb0 neisseria meningitidis (serogroup a)	1	W(R)XX(S)(Y)XG			

1	264: RDIKI	WRSVSYIG	LISYP		1	Q8YTB3	ck: 892	len: 369	! Q8YTB3 anabaena sp. (strain poc 7120).
	Q8ZPA8	ck: 2185	len: 372	! Q8ZPA8 salmonella typhimurium. alcohol dehy					
	309: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGSAFEG	VKGRT			309: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGSAFEG	ARGRT	
1	Q8ZNV6	ck: 6897	len: 261	! Q8ZNV6 salmonella typhimurium. abc superfam		Q8XCJ0	ck: 7492	len: 261	! Q8XCJ0 escherichia coli o157:h7. orf, h
	28: GSFVV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRRMSYFG	DTLAH			28: GSFVV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRRMSYFG	DTLAH	
1	Q8ZN10	ck: 8280	len: 199	! Q8ZN10 salmonella typhimurium. gifsy-1 prof		228: WRHRA	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRASAFTG	ICNDL	
	180: WRHRA	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRASAFTG	ICNDF			Q8X7X8	ck: 6047	len: 215	! Q8X7X8 escherichia coli o157:h7. hypoth
1	Q8ZMB6	ck: 4852	len: 1,181	! Q8ZMB6 salmonella typhimurium. exonuclease		65: VTPPQ	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRNKAFTG	LKDPE	
	901: LLYDS	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRVTSYSG	LQORG			Q8X6M9	ck: 2018	len: 1,180	! Q8X6M9 escherichia coli o157:h7. dna he
1	Q8ZH88	ck: 4557	len: 1,220	! Q8ZH88 yersinia pestis. exodeoxyribonucleas		901: LPGDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRVTSYSG	LQORG	
	904: KMQDY	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRVTSYSG	LQORG			Q8X5J4	ck: 7697	len: 369	! Q8X5J4 escherichia coli o157:h7. alcoho
1	Q8ZG44	ck: 7342	len: 766	! Q8ZG44 yersinia pestis. hypothetical protei		309: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGSAFEG	VKGRS	
	305: ADAFL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTISYFG	GKGVV			Q8UF80	ck: 5850	len: 298	! Q8UF80 agrobacterium tumefaciens (strai
1	Q8ZG18	ck: 6334	len: 377	! Q8ZG18 yersinia pestis. probable alcohol de		55: GCFVV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRMAYFG	DTMAH	
	310: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGSAFEG	VKGRS			Q8X3C3	ck: 8152	len: 227	! Q8X3C3 escherichia coli o157:h7. hypot
1	Q8Z6M8	ck: 4455	len: 122	! Q8Z6M8 salmonella typhi. putative pathogeni		35: DEIKY	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRVTSYSG	GKLPE	
	70: KALAN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKPAAFQG	IPOQL			Q9ACP5	ck: 2721	len: 1,039	! Q9ACP5 streptomyces coelicolor. putativ
1	Q8Z5W5	ck: 7341	len: 261	! Q8Z5W5 salmonella typhi. high-affinity zinc		603: GGAFV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKVRAYDG	SAYSA	
	28: GSFVV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRRMSYFG	DTLAH			Q9RRD0	ck: 3668	len: 206	! Q9RRD0 deinococcus radiodurans. 3-demet
1	Q8Z419	ck: 4255	len: 1,181	! Q8Z419 salmonella typhi. exonuclease v subu		111: VEETG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTASFDG	VVVG	
1	901: LLYDS	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRVTSYSG	LQORG			Q9ZP77	ck: 5059	len: 275	! Q9ZP77 rhizobium meliloti (sinorhizobii
							W(K,R)XX(S,A)(Y,F)XG		

29: GCFVI W(R)XX(A)(Y)XG DTMAR
 WRRMAIFG
 Q8YE09 ck: 9722 len: 370 ! Q8yeq9 bruceella melitensis. alcohol dehydro
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 310: VTGRV WKGTAFGG ARGRT
 Q8YDJ9 ck: 8470 len: 284 ! Q8yvj9 bruceella melitensis. high-affinity
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(Y)XG
 40: GCFII WRRMAIFG DTMAR
 Q8YLT4 ck: 3386 len: 368 ! Q8ylt4 ralstonia solanacearum (pseudomonas
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 308: VTGRE WKGSFAFG ARGRT
 Q8XTN7 ck: 4468 len: 368 ! Q8xtn7 ralstonia solanacearum (pseudomonas
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 308: VTGRE WKGSFAFG ARGRT
 Q8RHM9 ck: 4108 len: 454 ! Q8rhm9 fusobacterium nucleatum (subsp. nucl
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(S)(F)XG
 226: EIKSN WKSPSFSG VLEPE
 Q8REQ9 ck: 5980 len: 303 ! Q8req9 fusobacterium nucleatum (subsp. nucl
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(R)(Y)XG
 216: NIFLS WKSIAYAG FLSSG
 Q8R8C0 ck: 4513 len: 474 ! Q8r8c0 thermocanaerobacter tengcongensis. s-
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(Y)XG
 446: KQGGK WKLAAYDG VYNPV
 Q8R6S5 ck: 1574 len: 835 ! Q8r6s5 thermocanaerobacter tengcongensis. hy
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 608: YGSPT WKDIATIG TLNGA
 Q8PPQ7 ck: 6719 len: 697 ! Q8ppq7 xanthomonas axonopodis (pv. citri).
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 374: PSPDG WKKSFAVG APSFG
 Q8PPN3 ck: 2843 len: 368 ! Q8ppn3 xanthomonas axonopodis (pv. citri).
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 308: VTGRV WRGSFAFG VKGRS

Q8PPF2 ck: 6865 len: 369 ! Q8ppf2 xanthomonas axonopodis (pv. citri)
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 309: VTGRV WRGSFAFG VKGRT
 Q8PLX4 ck: 8710 len: 201 ! Q8plx4 xanthomonas axonopodis (pv. citri)
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 102: SPNGA WRNAAFQG YADHM
 Q8PHF1 ck: 333 len: 811 ! Q8phf1 xanthomonas axonopodis (pv. citri)
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 645: ELVAN WRQYAFGG DLLLT
 Q8PF27 ck: 1130 len: 423 ! Q8pf27 xanthomonas axonopodis (pv. citri)
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 146: GKAGR WRYPSEGG RTTIA
 Q8PA46 ck: 898 len: 205 ! Q8pa46 xanthomonas campestris (pv. campe
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 104: SPNGA WRNAAFQG YADHM
 Q8P691 ck: 8539 len: 811 ! Q8p691 xanthomonas campestris (pv. campe
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 645: ELVAN WRQYAFGG DLLLT
 Q8P5F2 ck: 6765 len: 369 ! Q8p5f2 xanthomonas campestris (pv. campe
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 309: VTGRV WRGSFAFG VKGRT
 Q8P568 ck: 4077 len: 368 ! Q8p568 xanthomonas campestris (pv. campe
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 308: VTGRV WRGSFAFG VKGRS
 Q8P541 ck: 7584 len: 697 ! Q8p541 xanthomonas campestris (pv. campe
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 374: PSDDG WKKSFAVG APAFG
 Q8P400 ck: 8382 len: 300 ! Q8p400 xanthomonas campestris (pv. campe
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 79: PTPAQ WKNVAFMG AALLL
 Q8NWJ8 ck: 6871 len: 412 ! Q8nwj8 staphylococcus aureus (strain mw2

1	35: YDFSP	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKNKSFNG	VINNT	Q8NSQ2	ck: 387	len: 275	! Q8nsq2 corynebacterium glutamicum (brevibac	1	Q9EWW0	ck: 2209	len: 266	! Q9eww0 streptomyces coelicolor. putativ
1	130: RAVIG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKDLAYAG	VIDSG	Q8NSJ2	ck: 6981	len: 301	! Q8nsj2 corynebacterium glutamicum (brevibac	1	239: VEDAA	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRSVAFSG	DRAEG	
1	254: AKRNW	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKLISFTG	IGSVV	Q8NQY7	ck: 9818	len: 530	! Q8nqy7 corynebacterium glutamicum (brevibac	1	306: ADAPL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTISYFG	GKGVV	
1	132: LREGG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKRSSFNG	VEIFG	Q9SZV4	ck: 1160	len: 330	! Q9szv4 streptomyces coelicolor. putative me	1	28: GSFVV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRRNSYFG	DTLAH	
1	31: KARLR	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRLAFAG	FVGVV	Q9RKB5	ck: 2102	len: 519	! Q9rkb5 streptomyces coelicolor. monooxygena	1	Q8G857	ck: 8641	len: 366	! Q8g857 bifidobacterium longum. permease
1	50: SYGGT	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRDNSYFG	CACDV	Q9RD13	ck: 4085	len: 660	! Q9rd13 streptomyces coelicolor. putative re	1	131: LMIFF	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRWFAG	LPSPS	
1	135: AALGE	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WEGPAYAG	FGARD	Q9RDQ2	ck: 398	len: 65	! Q9rdq2 streptomyces coelicolor. hypothetica	1	42: GCFII	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRRMAYFG	DTMAH	
1	12: LSNVE	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKRSSYSG	SNGGD	Q9RDI8	ck: 4543	len: 65	! Q9rdi8 streptomyces coelicolor. hypothetica	1	Q8FPV9	ck: 8770	len: 530	! Q8fpv9 corynebacterium efficiens. putat
1	7: SIMDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRRSSYSG	PGDGN	Q9RDI7	ck: 8793	len: 65	! Q9rdi7 streptomyces coelicolor. hypothetica	1	148: FWRVL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRFAYTG	QVIVA	
1	7: RRMDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKRSSYSG	PDDGN	Q9ADD0	ck: 1282	len: 243	! Q9add0 streptomyces coelicolor. hypothetica	1	309: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGSAFSG	VKGHS	
1	47: HRGHR	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKAASFGLG	VDEVY	Q8FGR4	ck: 7660	len: 261	! Q8fgr4 escherichia coli o6. high-affin.	1	Q8FPG1	ck: 8529	len: 369	! Q8fpg1 escherichia coli o6. alcohol del
									Q8FI60	ck: 9281	len: 247	! Q8fi60 escherichia coli o6. putative t
									228: WRHRA	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRASAFYG	IYNDL	

1 W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG DTLAH
28: GSFVY WRRMSYFG

Q8FG68 ck: 9474 len: 261 ! Q8fg68 escherichia coli o6. hypothetical pr
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
71: DEIKY WRTSYKG GEPPE

Q8FEB3 ck: 4588 len: 1,183 ! Q8feb3 escherichia coli o6. exodeoxyribonuc
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
904: LFGDN WRVTSYSG LQORG

Q8F6V8 ck: 7773 len: 523 ! Q8f6v8 leptospira interrogans. putative out
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
136: VTDFI WRGLSFSG EMANR

Q8F4M6 ck: 7773 len: 458 ! Q8f4m6 leptospira interrogans. inner membra
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
244: Q1SSN WKDPSFEG SFLPK

Q8F2V5 ck: 8001 len: 749 ! Q8f2v5 leptospira interrogans. ribonuclease
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
284: LESPI WRLASFEG VELKD

Q8EYP7 ck: 6380 len: 403 ! Q8eyp7 leptospira interrogans. argininosuc
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
149: TIIAP WRIWSFGG RSDLI

Q8EXC1 ck: 4765 len: 249 ! Q8exc1 leptospira interrogans. probable sug
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
211: KYMFS WKSPSYSG EPRKL

Q8BFC7 ck: 2100 len: 379 ! Q8efc7 shewanella oneidensis. alcohol dehyd
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
317: VTGRV WRGSAFEG VKGRS

Q8EF45 ck: 5478 len: 1,259 ! Q8ef45 shewanella oneidensis. exodeoxyribon
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
949: QVTRP WRVGSYSG LVKNA

Q8E800 ck: 4584 len: 376 ! Q8e800 shewanella oneidensis. zinc-binding
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG

314: VTGRV WKGSAFEG VKGRS

Q8DWE2 ck: 8469 len: 372 ! Q8dwe2 streptococcus mutans. putative al
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
312: VTGRV WRGSAFEG VKGKT

Q8DJN0 ck: 1651 len: 399 ! Q8djn0 synechococcus elongatus (thermosy
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
258: HNQVY WRNQSYYG FGGA

Q8DFQ3 ck: 1793 len: 265 ! Q8dfq3 vibrio vulnificus. abc-type mn2+/
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(Y)XG
32: GSFVW WRRMAYFG DTLAH

Q8DF77 ck: 391 len: 376 ! Q8df77 vibrio vulnificus. zn-dependent a
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
315: VTGRV WRGSAFEG VKGRS

Q8DBL8 ck: 5828 len: 1,206 ! Q8dbl8 vibrio vulnificus. atp-dependent
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
909: PIDRL WRITSYSG LVKQG

Q8D386 ck: 9466 len: 265 ! Q8d386 wigglesworthia brevipalpis. yebl
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
28: GSFLI WRKMSYFG DTLSH

Q8D070 ck: 2620 len: 379 ! Q8d070 yersinia pestis. alcohol dehydrog
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
312: VTGRV WRGSAFEG VKGRS

Q8CZY3 ck: 1453 len: 1,241 ! Q8czy3 yersinia pestis. dna helicase, at
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
925: KMQDY WRVTSYSG LQOSG

Q8CKW8 ck: 7447 len: 766 ! Q8ckw8 yersinia pestis. hypothetical. 3/
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
306: ADAFL WRTISYFG GKGVY

O58458 ck: 9251 len: 278 ! O58458 pyrococcus horikoshii. hypothetical
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
99: PERVG WRVSSYMG ISFON

```

1      Q97YM4 ck: 5165 len: 422 ! Q97ym4 sulfolobus solfataricus. n-acetylglu
      W(K,R)XX(S,A)(Y,F)XG
      W(R)XX(S)(Y)XG
415: YGVSE      WRGSSYLQ

1      Q8THB2 ck: 6769 len: 355 ! Q8thb2 methanosarcina acetivorans. iron(III)
      W(K,R)XX(S,A)(Y,F)XG
      W(R)XX(A)(F)XG
24: PFCLF      WRSIAFIG      FLLLP

1      Q8Q0X8 ck: 4651 len: 535 ! Q8q0x8 methanosarcina mazei (methanosarcina
      W(K,R)XX(S,A)(Y,F)XG
      W(K)XX(A)(Y)XG
469: QAIEN      WKLAAYDG      NTGFG

```

Databases searched:

SWISS-PROT, Release 41.1, Released on 6Jun2003, Formatted on 9Jun2003
 SPTREMBL, Release 23.0, Released on 4Mar2003, Formatted on 7Mar2003

Total finds: 336
 Total length: 305,079,309
 Total sequences: 958,388
 CPU time: 07:22.84

```

! FINDPATTERNS on pir:* allowing 0 mismatches
! 1 W(K,R)XX(S,A)(Y,F)XG August 15, 2003 10:46 ..
DEHUC2 ck: 9148 len: 374 ! alcohol dehydrogenase (EC 1.1.1.1) 5 [valid
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
314: VTGRT WKGTAFFG WKSVE

A33419 ck: 6584 len: 373 ! alcohol dehydrogenase (EC 1.1.1.1) class II
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
313: VTGRT WKGTAFFG WKSVE

DERTA ck: 6609 len: 373 ! alcohol dehydrogenase (EC 1.1.1.1) 2 - rat
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
313: VTGRT WKGTAFFG WKSVE

A56643 ck: 1156 len: 374 ! alcohol dehydrogenase (EC 1.1.1.1) 2 - mous
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
314: VTGRT WKGTAFFG WKSVE

S68061 ck: 5507 len: 373 ! alcohol dehydrogenase (EC 1.1.1.1) class II
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
313: VTGRT WKGTAFFG WKSVE

JC4967 ck: 6805 len: 376 ! alcohol dehydrogenase (EC 1.1.1.1) class II
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
316: VTGRV WKGTAFFG WKSVE

S51187 ck: 4299 len: 376 ! alcohol dehydrogenase (EC 1.1.1.1) class II
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
316: VTGRT WKGTAFFG WKSVE

A49662 ck: 6276 len: 378 ! alcohol dehydrogenase (EC 1.1.1.1) class II
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
316: VTGRV WKGTAFFG FKSRD

S51357 ck: 1032 len: 379 ! alcohol dehydrogenase (EC 1.1.1.1) Fdh - fr
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
319: VTGRV WKGSAFAG WKSVS

S71244 ck: 5257 len: 379 ! alcohol dehydrogenase (EC 1.1.1.1) class II
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
317: VTGRV WKGTAFFG FKSRD

D64763 ck: 7661 len: 369 ! alcohol dehydrogenase (EC 1.1.1.1) C - E
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
309: VTGRV WKGSAFAG VKGRS

S57525 ck: 7402 len: 369 ! alcohol dehydrogenase (EC 1.1.1.1) C - E
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
309: VTGRV WKGSAFAG VKGRT

H64052 ck: 2043 len: 378 ! alcohol dehydrogenase (EC 1.1.1.1) H101E
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
318: VTGRV WKGSAFAG VKGRS

S31140 ck: 583 len: 386 ! alcohol dehydrogenase (EC 1.1.1.1) SFAL
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
322: VTGRV WKGSAFAG IKGRS

JN0447 ck: 8341 len: 381 ! alcohol dehydrogenase (EC 1.1.1.1) FDH1
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
320: VTGRT WKGAAFFG VKGRS

B54075 ck: 7421 len: 663 ! arachidonate 12-lipoxygenase (EC 1.13.11)
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
223: RVRRS WKEDAFTG YQFLN

S30051 ck: 6102 len: 663 ! arachidonate 12-lipoxygenase (EC 1.13.11)
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
223: RVRRS WKEDAFTG YQFLN

I38344 ck: 431 len: 26,926 ! titin, cardiac muscle [validated] - hum
1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
18,126: NVTLK WKKPAYDG GSKIT

NCECX5 ck: 1854 len: 1,180 ! exodeoxyribonuclease V (EC 3.1.11.5) 13:
1 W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
901: LPGDN WRTVSYSG LQQRG

A35742 ck: 2790 len: 513 ! aqualysin (EC 3.4.21.-) I precursor - nt
1 W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(Y)XG
495: AGYIL WKTIAYSG SGMYE

S33643 ck: 8565 len: 733 ! transforming protein B-myo - African cle

```

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
705: PMTAA WKVAFGG SQDQM
QOCV2 ck: 3030 len: 159 ! aphid transmission protein - cauliflower mo
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
61: SLIGI WKINSIFG LSKDP
TJBPCL ck: 9747 len: 199 ! tail assembly protein K - phage lambda
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
180: WRHRA WRSAFTG IYNDL
S22923 ck: 2030 len: 308 ! ubiquinol-cytochrome-c reductase (EC 1.10.2
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
103: GYVLP WRQMSFWG ATVIT
T03289 ck: 4477 len: 381 ! formaldehyde dehydrogenase (glutathione) (E
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
319: VTGRV WKGTAFGG FKSRT
T04164 ck: 4843 len: 381 ! formaldehyde dehydrogenase (glutathione) (E
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
319: VTGRV WKGTAFGG FKSRS
T40965 ck: 1472 len: 380 ! alcohol dehydrogenase (EC 1.1.1.1) class II
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
319: VTGRV WRGCAFGG VKGRS
F81097 ck: 788 len: 378 ! probable alcohol dehydrogenase (EC 1.1.1.1)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
318: VTGRV WKGSAFGG VKGRS
B83191 ck: 6470 len: 370 ! alcohol dehydrogenase (EC 1.1.1.1) [similar
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
310: VTGRV WRGSAFGG VKGRS
JC7759 ck: 5519 len: 376 ! alcohol dehydrogenase (EC 1.1.1.1) 3 - zebu
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
316: VTGRV WKGTAFGG WKSVE
AB3479 ck: 9722 len: 370 ! alcohol dehydrogenase (EC 1.1.1.1) [importe
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG

310: VTGRV WKGTAFGG ARGRT
AB0183 ck: 6334 len: 377 ! alcohol dehydrogenase (EC 1.1.1.1) [simi
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
310: VTGRV WRGSAFGG VKGRS
C87561 ck: 8539 len: 369 ! alcohol dehydrogenase (EC 1.1.1.1) [simi
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
309: VTGRV WKGTAFGG ARGRT
C90680 ck: 7697 len: 369 ! alcohol dehydrogenase (EC 1.1.1.1) [simi
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
309: VTGRV WKGSAFGG VKGRS
G85530 ck: 7697 len: 369 ! alcohol dehydrogenase (EC 1.1.1.1) [simi
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
309: VTGRV WKGSAFGG VKGRS
AC2157 ck: 892 len: 369 ! alcohol dehydrogenase (EC 1.1.1.1) [simi
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
309: VTGRV WKGSAFGG ARGRT
B83850 ck: 2788 len: 540 ! D-3-phosphoglycerate dehydrogenase BH160
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
141: IKAGE WKRAFGG TELRG
G70854 ck: 1055 len: 528 ! probable sera protein - Mycobacterium tu
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
129: LREHT WKRSFSG TEIFG
T45418 ck: 421 len: 528 ! phosphoglycerate dehydrogenase [imported
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
129: LRAHI WKRSFSG TEIFG
I52462 ck: 6502 len: 663 ! arachidonate 12-lipoxygenase (EC 1.13.11
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
223: RVRNS WKEDAFGG YQFLN
S31959 ck: 845 len: 396 ! acyl-[acyl-carrier-protein] desaturase (E
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
340: FLVGR WKVDAFTG LSGEG

1	S20901	ck: 1434	len: 6,805	! titin - rabbit (fragment)	415: YGVSE	W(R)xx(S)(Y)xg WRGSSYLK		
	6,485: NVTLK	W(K,R)XX(S,A)(Y,F)xg W(K)xx(A)(Y)xg WKKPAYDS	GSKIT					
1	E91088	ck: 2018	len: 1,180	! DNA helicase RecB [imported] - Escherichia	D70475	ck: 5280	len: 392	! conserved hypothetical protein aq_2044
	901: LFGDN	W(K,R)XX(S,A)(Y,F)xg W(R)xx(S)(Y)xg WRVTSYSG	LQQRG		127: YVSVN	W(K,R)XX(S,A)(Y,F)xg W(K)xx(S)(F)xg WKRSSFEG	KDFED	
1	B82091	ck: 9551	len: 1,208	! exodeoxyribonuclease V, 135 kDa chain VC232	E72167	ck: 405	len: 146	! A32L protein - variola minor virus (str
	906: AIDRR	W(K,R)XX(S,A)(Y,F)xg W(R)xx(S)(Y)xg WRVTSYSG	LVMQS		73: DVKQK	W(K,R)XX(S,A)(Y,F)xg W(R)xx(A)(Y)xg WRCVAYPG	NGFVS	
1	G85933	ck: 2018	len: 1,180	! DNA helicase RecB [imported] - Escherichia	JQ1834	ck: 405	len: 146	! 3L protein - variola major virus
	901: LFGDN	W(K,R)XX(S,A)(Y,F)xg W(R)xx(S)(Y)xg WRVTSYSG	LQQRG		73: DVKQK	W(K,R)XX(S,A)(Y,F)xg W(R)xx(A)(Y)xg WRCVAYPG	NGFVS	
1	AD0125	ck: 4557	len: 1,220	! exodeoxyribonuclease V (EC 3.1.11.5) beta c	S46858	ck: 405	len: 146	! A31L protein - variola virus
	904: KMQDY	W(K,R)XX(S,A)(Y,F)xg W(R)xx(S)(Y)xg WRVTSYSG	LQQRG		73: DVKQK	W(K,R)XX(S,A)(Y,F)xg W(R)xx(A)(Y)xg WRCVAYPG	NGFVS	
1	AB0865	ck: 4255	len: 1,181	! exonuclease V chain [imported] - Salmonella	H90733	ck: 8757	len: 199	! probable tail assembly protein [imported]
	901: LLYDS	W(K,R)XX(S,A)(Y,F)xg W(R)xx(S)(Y)xg WRVTSYSG	LQQRG		180: WRHRA	W(K,R)XX(S,A)(Y,F)xg W(R)xx(A)(F)xg WRASAFYG	ICNDL	
1	T16557	ck: 2841	len: 551	! hypothetical protein K04E7.3 - Caenorhabdit	F90834	ck: 2142	len: 247	! tail assembly protein [imported] - Esch
	376: RADYS	W(K,R)XX(S,A)(Y,F)xg W(R)xx(S)(Y)xg WRNISYSG	SSDCY		228: WRHRA	W(K,R)XX(S,A)(Y,F)xg W(R)xx(A)(F)xg WRASAFYG	ICNDL	
1	S41178	ck: 957	len: 159	! gene 36 protein - phage SP1	B85584	ck: 3865	len: 224	! probable tail component of prophage CP-
	77: GVSQS	W(K,R)XX(S,A)(Y,F)xg W(R)xx(A)(F)xg WRTGAFEG	QDGKR		205: WRHRA	W(K,R)XX(S,A)(Y,F)xg W(R)xx(A)(F)xg WRASAFYG	ICNDL	
1	H64449	ck: 6435	len: 70	! ribosomal protein L24E - Methanococcus jann	A98950	ck: 7492	len: 261	! hypothetical protein ECs2569 [imported]
	4: MPE	W(K,R)XX(S,A)(Y,F)xg W(R)xx(S)(F)xg WRTCSFCG	YEIEP		28: GSFVV	W(K,R)XX(S,A)(Y,F)xg W(R)xx(S)(Y)xg WRRMSYFG	DTLAH	
1	S23164	ck: 5378	len: 65	! light-harvesting protein alpha chain - Ecto	F82120	ck: 3098	len: 260	! zinc ABC transporter, permease protein 1
	58: ENPGI	W(K,R)XX(S,A)(Y,F)xg W(K)xx(S)(Y)xg WRTSYDG			28: GSFVV	W(K,R)XX(S,A)(Y,F)xg W(R)xx(A)(Y)xg WRRMAYFG	DTLAH	
1	A99285	ck: 5165	len: 422	! hypothetical protein nodC-like [imported] -	F82959	ck: 5958	len: 262	! permease of ABC zinc transporter ZnuB p
		W(K,R)XX(S,A)(Y,F)xg			28: GSFVV	W(K,R)XX(S,A)(Y,F)xg W(R)xx(A)(Y)xg WRRMAYFG	DTLSH	

1 A64066 ck: 1426 len: 261 ! probable membrane protein HI0407 - Haemophilus influenzae
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(F)XG
28: GNFVV WRKMAIFG DTLSE

1 E85798 ck: 7492 len: 261 ! hypothetical protein yebI [imported] - Escherichia coli
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(Y)XG
28: GSFVV WRMSYFVG DTLAH

1 AC0251 ck: 7586 len: 261 ! high-affinity zinc uptake system membrane protein
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(Y)XG
28: GSFVV WRMSYFVG DTLAH

1 AF3531 ck: 8470 len: 284 ! high-affinity zinc uptake system membrane protein
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(Y)XG
40: GCFII WRMAIFG DTMAH

1 C64948 ck: 7831 len: 261 ! probable membrane protein yebI - Escherichia coli
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(Y)XG
28: GSFVV WRMSYFVG DTLAH

1 AE2763 ck: 9687 len: 272 ! hypothetical protein znuB [imported] - Agrobacterium tumefaciens
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(Y)XG
29: GCFVV WRMAIFG DTMAH

1 AI0742 ck: 7341 len: 261 ! high-affinity zinc uptake system membrane protein
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(Y)XG
28: GSFVV WRMSYFVG DTLAH

1 E97544 ck: 5950 len: 298 ! permease of ABC zinc transporter znuB (PA55)
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(Y)XG
55: GCFVV WRMAIFG DTMAH

1 S65290 ck: 8779 len: 475 ! clathrin-associated protein complex medium
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(S)(F)XG
382: KSAIL WKIRSFPG GXEYS

1 C96633 ck: 2687 len: 428 ! probable Serine/Threonine protein kinase F8
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(F)XG
346: KDALV WKIKSFPG NKEYM

1 D70786 ck: 6149 len: 379 ! probable gcvT protein - Mycobacterium tuberculosis
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(Y)XG

1 W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(F)XG
275: GWAVG WRKDAFFG RAALL

1 T44888 ck: 3832 len: 367 ! probable aminomethyltransferase (EC 2.1.1.)
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(F)XG
263: GWAIG WRKDAFLG RDALL

1 A39484 ck: 2015 len: 280 ! androgen-withdrawal apoptosis protein RV
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(F)XG
29: CALPM NEVSAFIG SSIIT

1 T32510 ck: 5742 len: 246 ! hypothetical protein C44B12.3 - Caenorhabditis elegans
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(S)(Y)XG
87: VVKMT WRKQSYDG VLAFV

1 T00479 ck: 1756 len: 303 ! probable phosphatidylinositol-glycan synthase
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(Y)XG
15: IFGPK WRKVAYGG MQIGY

1 A56152 ck: 5845 len: 213 ! major 25k outer membrane protein precursors
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(F)XG
74: IKPDD WKAGAFAG WNFOQ

1 AC3408 ck: 5809 len: 213 ! 25k outer-membrane immunogenic protein P
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(F)XG
74: IKPDD WKAGAFAG WNFOQ

1 G85506 ck: 391 len: 198 ! hypothetical protein 20246 [imported] - Escherichia coli
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(F)XG
48: VTTPQ WRNKAFNG LKDPE

1 F90655 ck: 6047 len: 215 ! hypothetical protein ECs0214 [imported]
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(F)XG
65: VTTPQ WRNKAFNG LKDPE

1 G90978 ck: 8152 len: 227 ! hypothetical protein ECs2799 [imported]
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(Y)XG
35: DEIKY WRNYSYNG GKLPE

1 AF0118 ck: 8384 len: 765 ! probable kinase YPO0966 [imported] - Yersinia enterocolitica
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(Y)XG
306: ADAPL WRTISYNG GKGVV

```

1 AG0179 ck: 7342 len: 766 ! hypothetical protein YP01473 [imported] - y
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
402: ERFEI WKQSFIG KLHNF

306: ADAPL WRTISYFG GKGVV

1 AG0697 ck: 4455 len: 122 ! probable pathogenicity island protein STY17
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
70: KALAN WKPAPFG IPQRL

1 F87390 ck: 9243 len: 613 ! TonB-dependent receptor, probable [imported]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
370: KGDW WRAAYAG FRPPT

1 S60618 ck: 760 len: 441 ! 4-alpha-glucanotransferase (EC 2.4.1.25) -
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(Y)XG
348: EQTF WKWPAYNG PPSGI

1 F83354 ck: 8784 len: 538 ! probable sulfatase PA2333 [imported] - Pseu
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(Y)XG
455: VRSE WKYIAYDG FRAQL

1 S78561 ck: 548 len: 937 ! CS3 pilin synthesis protein, 104K - Escheri
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
595: SIETD WRGAFIG YLSPY

1 H75258 ck: 3668 len: 206 ! probable 3-demethylubiquinone-9 3-methyltra
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
111: VEETG WRTASFDG VVVG

1 T34972 ck: 1160 len: 330 ! probable membrane protein - Streptomyces co
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
31: KARLR WLLAFAG FVGVV

1 T06088 ck: 7947 len: 700 ! hypothetical protein T9A14.170 - Arabidopsi
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
645: PQOIN WRPPSYFG RDDLE

1 T02995 ck: 2338 len: 530 ! unspecific monooxygenase (EC 1.14.14.1) - o
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
254: RGCSP WRPDAGV GNDYY

1 T00208 ck: 8858 len: 777 ! transposase-like protein - fungus (Fusarium

```

Databases searched:
 NBRF, Release 76.1, Released on 12May2003, Formatted on 10Jun2003

Total finds: 98
 Total length: 96,168,682
 Total sequences: 283,308
 CPU time: 02:11.82

! FINDPATTERNS on geneseqp: * allowing 0 mismatches

1 1 W(K,R)XX(S,A)(Y,F)XG August 15, 2003 10:54 ..

1 AAR04585 ck: 2828 len: 513 ! Aar04585 Aquaricine I. 3/2003
W(K,R)XX(S,A)(Y,F)XG
495: AGYLL WRIYAYSG SGMYE

1 AAR13181 ck: 2828 len: 513 ! Aar13181 T.aquaticus Aquaricin I. 10/1991
W(K,R)XX(S,A)(Y,F)XG
495: AGYLL WRIYAYSG SGMYE

1 AAR67653 ck: 2828 len: 513 ! Aar67653 Aqualysin I. 8/1995
W(K,R)XX(S,A)(Y,F)XG
495: AGYLL WRIYAYSG SGMYE

1 AAR97244 ck: 1074 len: 4,473 ! Aar97244 Virulence gene cluster polypeptide
W(K,R)XX(S,A)(Y,F)XG
1,144: KALAN WKPAAFQG IPQRL

1 AAW82254 ck: 7849 len: 934 ! Aaw82254 JP10248575 Seq ID 4. 7/1999
W(K,R)XX(S,A)(Y,F)XG
79: EDGTV WRSRAVHG KLGKY

1 AAY41726 ck: 2066 len: 211 ! Aay41726 Human PRO944 protein sequence. 12/
W(K,R)XX(S,A)(Y,F)XG
30: TALPQ WRIYSYAG DNIVT

1 AAY30337 ck: 9802 len: 1,091 ! Aay30337 Protein encoded by the human DLC-1
W(K,R)XX(S,A)(Y,F)XG
401: NGSVN WRTGSPFHG PGHIS

1 AAY31650 ck: 884 len: 530 ! Aay31650 Brevibacterium flavum mutant D-3-ph
W(K,R)XX(S,A)(Y,F)XG
132: LREGE WKRSSFNG VEIFG

1 AAY31651 ck: 6831 len: 345 ! Aay31651 Corynebacterium glutamicum D-3-ph
W(K,R)XX(S,A)(Y,F)XG
132: LREGE WKRSSFNG VEIFG

1 AAY31649 ck: 644 len: 530 ! Aay31649 Brevibacterium flavum wild-type D-
W(K,R)XX(S,A)(Y,F)XG
132: LREGE WKRSSFNG VEIFG

1 AAY38430 ck: 5842 len: 212 ! Aay38430 Human secreted protein. 9/1999
W(K,R)XX(S,A)(Y,F)XG
30: TALPQ WRIYSYAG DNIVT

1 AAY38421 ck: 4249 len: 71 ! Aay38421 Human secreted protein encoded
W(K,R)XX(S,A)(Y,F)XG
30: TALPQ WRIYSYAG DNIVT

1 AAY36134 ck: 9084 len: 230 ! Aay36134 Human secreted protein #6. 9/1
W(K,R)XX(S,A)(Y,F)XG
30: MLLPS WKTSSYVG ASIVT

1 AAY36181 ck: 8102 len: 230 ! Aay36181 Human secreted protein #53. 9/
W(K,R)XX(S,A)(Y,F)XG
30: MLLPS WKTSSYVG ASIVT

1 AAY06346 ck: 8232 len: 84 ! Aay06346 EGIII-like cellulase (partial
W(K,R)XX(S,A)(Y,F)XG
54: SIDST WKWKSYSY SNIVA

1 AAY22645 ck: 6831 len: 345 ! Aay22645 3-PGDH protein, also known as
W(K,R)XX(S,A)(Y,F)XG
132: LREGE WKRSSFNG VEIFG

1 AAY22646 ck: 644 len: 530 ! Aay22646 Wild type 3-PGDH protein, also
W(K,R)XX(S,A)(Y,F)XG
132: LREGE WKRSSFNG VEIFG

1 AAY22647 ck: 884 len: 530 ! Aay22647 Mutant 3-PGDH protein, also kn
W(K,R)XX(S,A)(Y,F)XG
132: LREGE WKRSSFNG VEIFG

1 AAY34536 ck: 9824 len: 951 ! Aay34536 Porphyromonas gingivalis prote
W(K,R)XX(S,A)(Y,F)XG
576: GMFKP WKPFPSFG NLIMI

1 AAY34403 ck: 6251 len: 953 ! Aay34403 Porphyromonas gingivalis prote
W(K,R)XX(S,A)(Y,F)XG
578: GMFKP WKPFPSFG NLIMI

1 AAY13939 ck: 1473 len: 220 ! Aay13939 Human transmembrane protein, H
W(K,R)XX(S,A)(Y,F)XG
132: LREGE WKRSSFNG VEIFG

1	29: CALPM	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVSFAFG	SNIT	1	25: RDSGR	WREVSFFG	ETERA
1	AAV12226	ck: 4096 len: 114	! Aay12226 Human 5' EST secreted protein SEQ	1	AAB53886	ck: 6120 len: 75	! Aab53886 Human colon cancer antigen prot
1	30: MLPLS	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKTSSVVG	ASIVT	60: GXIPP	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKLXSFLG	KRXKX	
1	AAV12227	ck: 8204 len: 55	! Aay12227 Human 5' EST secreted protein SEQ	1	AAB54052	ck: 8653 len: 155	! Aab54052 Human pancreatic cancer antigen
1	30: MLPLS	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKTSSVVG	ASIVT	55: MLPLS	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKTSSVVG	ASIVT	
1	AAV04143	ck: 2066 len: 211	! Aay04143 Human Tango-73 protein. 6/1999	1	AAB52100	ck: 5242 len: 208	! Aab52100 Gene 48 human secreted protein
1	30: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRIYSYAG	DNIVT	30: CALPM	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVTAFIG	SNIVT	
1	AAW99653	ck: 2066 len: 211	! Aaw99653 Human senescence factor p23 protein	1	AAB12138	ck: 3847 len: 497	! Aab12138 Hydrophobic domain protein isol
1	30: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRIYSYAG	DNIVT	4: MAL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRGSAYAG	FLALA	
1	AAW88747	ck: 4456 len: 206	! Aaw88747 Secreted protein encoded by gene 4	1	AAB28673	ck: 2647 len: 434	! Aab28673 Human carbohydrate-modifying en
1	192: NLLGG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKYSFSG	FLIFF	365: EMGLC	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKEVAILG	NEVSD	
1	AAW88629	ck: 6151 len: 202	! Aaw88629 Secreted protein encoded by gene 9	1	AAB43133	ck: 6154 len: 209	! Aab43133 Human ORFX ORF2897 polypeptide
1	22: CALPM	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVTAFIG	SNIVT	30: CALPM	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVTAFIG	SNIVT	
1	AAW86307	ck: 1982 len: 210	! Aaw86307 Kidney injury associated molecule	1	AAB43165	ck: 2062 len: 434	! Aab43165 Human ORFX ORF2929 polypeptide
1	30: CSLPM	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVTAFIG	SNIVT	365: EMGLC	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKEVAILG	NEVSD	
1	AAW72976	ck: 2905 len: 376	! Aaw72976 Rhodobacter sphaeroides adhI forma	1	AAB43405	ck: 6708 len: 383	! Aab43405 Human cancer associated protein
1	316: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGSAFGG	ARGRT	323: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGTAFFG	WKSVE	
1	AAW76765	ck: 2905 len: 376	! Aaw76765 R. sphaeroides AdhI class III alo	1	AAB43931	ck: 5762 len: 448	! Aab43931 Human cancer associated protein
1	316: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGSAFGG	ARGRT	379: EMGLC	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKEVAILG	NEVSD	
1	AAB53841	ck: 4392 len: 120	! Aab53841 Human colon cancer antigen protein	1	AAB44282	ck: 2066 len: 211	! Aab44282 Human PRO944 (UNQ481) protein s
1		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG		30: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRIYSYAG	DNIVT	

1	AAG44736	ck: 1284	len: 255	! Aag44736 Zea mays protein fragment SEQ ID N	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	E
	193: VTGRV	W(K)XX(A)(F)XG WKGTFAGG	FKSRT		2:	K	
1	AAG45941	ck: 3992	len: 395	! Aag45941 Arabidopsis thaliana protein fragm	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGTFAGG	K
	333: VTGRV	W(K)XX(A)(F)XG WKGTFAGG	FKSRT		2:	D	
1	AAG45942	ck: 5269	len: 379	! Aag45942 Arabidopsis thaliana protein fragm	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGTFAGG	K
	317: VTGRV	W(K)XX(A)(F)XG WKGTFAGG	FKSRT		2:	E	
1	AAG45943	ck: 2760	len: 255	! Aag45943 Arabidopsis thaliana protein fragm	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGTFAGG	K
	193: VTGRV	W(K)XX(A)(F)XG WKGTFAGG	FKSRT		1:		
1	AAB06419	ck: 2784	len: 8	! Aab06419 Claudin-1 cell adhesion recognitio	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	
	1:				2:	C	
1	AAB06420	ck: 3486	len: 9	! Aab06420 Claudin-1 cell adhesion recognitio	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	D
	1:				2:	K	
1	AAB06426	ck: 2798	len: 8	! Aab06426 Claudin-1 cell adhesion recognitio	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKIYSYAG	E
	1:				2:	K	
1	AAB06427	ck: 3500	len: 9	! Aab06427 Claudin-1 cell adhesion recognitio	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKIYSYAG	K
	1:				2:	D	
1	AAB06436	ck: 4153	len: 10	! Aab06436 Claudin-1 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	K
	2:	C			2:	E	
1	AAB06445	ck: 4171	len: 10	! Aab06445 Claudin-1 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	K
	2:	K			1:		
1	AAB06454	ck: 4181	len: 10	! Aab06454 Claudin-1 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTSSYVG	
	2:	K			1:		
1	AAB06463	ck: 4234	len: 10	! Aab06463 Claudin-1 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	E
	2:	D			2:	K	
1	AAB06472	ck: 4235	len: 10	! Aab06472 Claudin-1 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	K
	2:	E			2:	K	
1	AAB06479	ck: 2784	len: 8	! Aab06479 Claudin-1 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	
	1:				1:		
1	AAB06485	ck: 4174	len: 10	! Aab06485 Claudin-1 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKIYSYAG	
	2:	C			2:	C	
1	AAB06491	ck: 4192	len: 10	! Aab06491 Claudin-1 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKIYSYAG	D
	2:	K			2:	K	
1	AAB06497	ck: 4202	len: 10	! Aab06497 Claudin-1 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKIYSYAG	E
	2:	K			2:	K	
1	AAB06503	ck: 4255	len: 10	! Aab06503 Claudin-1 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKIYSYAG	K
	2:	D			2:	D	
1	AAB06509	ck: 4256	len: 10	! Aab06509 Claudin-1 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKIYSYAG	K
	2:	E			2:	E	
1	AAB06512	ck: 2798	len: 8	! Aab06512 Claudin-1 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WKIYSYAG	
	1:				1:		
1	AAB06521	ck: 2954	len: 8	! Aab06521 Claudin-2 cell adhesion recogni	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTSSYVG	
	2:	K			1:		

1	AAB06530	ck: 4356 len: 10	! Aab06530 Claudin-2 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	C	2: C WRTSSYVG	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	K
1	AAB06539	ck: 4374 len: 10	! Aab06539 Claudin-2 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(S)(Y)XG	D	2: K WRTSSYVG	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	K
1	AAB06548	ck: 4384 len: 10	! Aab06548 Claudin-2 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(S)(Y)XG	E	2: K WRTSSYVG	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	K
1	AAB06557	ck: 4437 len: 10	! Aab06557 Claudin-2 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(S)(Y)XG	K	2: D WRTSSYVG	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	K
1	AAB06566	ck: 4438 len: 10	! Aab06566 Claudin-2 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(S)(Y)XG	K	2: E WRTSSYVG	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	K
1	AAB06574	ck: 2954 len: 8	! Aab06574 Claudin-2 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(S)(Y)XG		1: WRTSSYVG	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06583	ck: 2665 len: 8	! Aab06583 Claudin-3 cell adhesion recognit	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG		1: WRTSSYVG	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06592	ck: 4019 len: 10	! Aab06592 Claudin-3 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG	C	2: C WRTSSYVG	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	K
1	AAB06601	ck: 4037 len: 10	! Aab06601 Claudin-3 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG	D	2: K WRTSSYVG	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	K
1	AAB06610	ck: 4047 len: 10	! Aab06610 Claudin-3 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG	E	2: K WRTSSYVG	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	K
1	AAB06619	ck: 4100 len: 10	! Aab06619 Claudin-3 cyclic cell adhesion rec	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06628	ck: 4101 len: 10	! Aab06628 Claudin-3 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06636	ck: 2665 len: 8	! Aab06636 Claudin-3 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06645	ck: 2669 len: 8	! Aab06645 Claudin-4 cell adhesion recogn	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06654	ck: 4024 len: 10	! Aab06654 Claudin-4 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06663	ck: 4042 len: 10	! Aab06663 Claudin-4 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06672	ck: 4052 len: 10	! Aab06672 Claudin-4 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06681	ck: 4105 len: 10	! Aab06681 Claudin-4 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06690	ck: 4106 len: 10	! Aab06690 Claudin-4 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06698	ck: 2669 len: 8	! Aab06698 Claudin-4 cyclic cell adhesion	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	
1	AAB06764	ck: 2655 len: 8	! Aab06764 Claudin-6/9 cell adhesion reco	1	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG			W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG W(R)xx(S)(Y)XG	

```

1:      WKVTAFIG
AAB06772  ck: 4003  len: 10  ! Aab06772 Claudin-6/9 cyclic cell adhesion r
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(A)(F)XG
2:      C      WKVTAFIG      C
AAB06781  ck: 4021  len: 10  ! Aab06781 Claudin-6/9 cyclic cell adhesion r
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(A)(F)XG
2:      K      WKVTAFIG      D
AAB06789  ck: 4031  len: 10  ! Aab06789 Claudin-6/9 cyclic cell adhesion r
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(A)(F)XG
2:      K      WKVTAFIG      E
AAB06798  ck: 4084  len: 10  ! Aab06798 Claudin-6/9 cyclic cell adhesion r
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(A)(F)XG
2:      D      WKVTAFIG      K
AAB06806  ck: 4085  len: 10  ! Aab06806 Claudin-6/9 cyclic cell adhesion r
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(A)(F)XG
2:      E      WKVTAFIG      K
AAB06814  ck: 2655  len: 8    ! Aab06814 Claudin-6/9 cyclic cell adhesion r
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(A)(F)XG
1:
AAB06911  ck: 4176  len: 10  ! Aab06911 Claudin cell adhesion recognition
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(S)(Y)XG
1:      WKLIISYAG      DN
AAB06916  ck: 2655  len: 8    ! Aab06916 Claudin cell adhesion recognition
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(A)(F)XG
1:      WKVTAFIG
AAY99378  ck: 8102  len: 230  ! Aay99378 Human PRO1356 (UNQ705) amino acid
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(S)(Y)XG
30:      MLLPS      WKTSSYVG      ASIIVT
AAY99420  ck: 828   len: 205  ! Aay99420 Human PRO1486 (UNQ755) amino acid
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(S)(F)XG
192:      NLLGG      WKYSSFSG      FLIFP

```

```

AAY99434  ck: 4773  len: 220  ! Aay99434 Human PRO1488 (UNQ757) amino ac
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(A)(F)XG
30:      CALPM      WKVTAFIG      NSIIV
AAY84609  ck: 8102  len: 230  ! Aay84609 A human membrane associated org
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(S)(Y)XG
30:      MLLPS      WKTSSYVG      ASIIVT
AAY51675  ck: 2898  len: 211  ! Aay51675 Murine clodin 1 protein. 6/2000
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(S)(Y)XG
30:      TALPQ      WKIISYAG      DNIIVT
AAY51676  ck: 6271  len: 230  ! Aay51676 Murine clodin 2 protein. 6/2000
      W(K,R)XX(S,A)(Y,F)XG
      W(R)xx(S)(Y)XG
30:      MLLPN      WKTSSYVG      ASIIVT
AAY51679  ck: 7312  len: 219  ! Aay51679 Murine clodin 3 protein. 6/2000
      W(K,R)XX(S,A)(Y,F)XG
      W(R)xx(A)(F)XG
29:      CALPM      WNVSAFVG      SSIIT
AAY51681  ck: 3382  len: 219  ! Aay51681 Murine clodin 6 protein. 6/2000
      W(K,R)XX(S,A)(Y,F)XG
      W(K)xx(A)(F)XG
30:      CALPM      WKVTAFIG      NSIIV
AAY44794  ck: 5085  len: 379  ! Aay44794 P. pastoris Formaldehyde Dehydr
      W(K,R)XX(S,A)(Y,F)XG
      W(R)xx(A)(F)XG
318:      VTGRV      WEGCAFGG      IKGR
AAY44795  ck: 7004  len: 392  ! Aay44795 P. pastoris Formaldehyde Dehydr
      W(K,R)XX(S,A)(Y,F)XG
      W(R)xx(A)(F)XG
318:      VTGRV      WEGCAFGG      IKGR
AAY68679  ck: 2066  len: 211  ! Aay68679 A human molecule associated wit
      W(K,R)XX(S,A)(Y,F)XG
      W(R)xx(S)(Y)XG
30:      TALPQ      WKIISYAG      DNIIVT
AAY76130  ck: 5842  len: 212  ! Aay76130 Human secreted protein encoded
      W(K,R)XX(S,A)(Y,F)XG
      W(R)xx(S)(Y)XG
30:      TALPQ      WKIISYAG      DNIIVT
AAY53588  ck: 8341  len: 381  ! Aay53588 Hepatitis B virus surface antig
      W(K,R)XX(S,A)(Y,F)XG

```

1	320: VTGRT	W(K)xx(A)(F)XG WKGARFGG	VKGRS	ABP41791	ck: 8440 len: 268	! Abp41791 Human ovarian antigen HSYBI49,
1	ABP53938	ck: 3468 len: 9	! Abp53938 VEGFR-3 binding peptide SEQ ID NO:	87: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(R)xx(S)(Y)XG	DNIVT
2:	R	WRGNAIFG		ABP41820	ck: 8565 len: 210	! Abp41820 Human ovarian antigen HOGCS42,
1	ABG95890	ck: 8102 len: 230	! Abg95890 Human secreted/transmembrane prote	31: CALPM	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG	SNIVT
30: MLLPS		W(K)xx(S)(Y)XG WKTSSYVG	ASIVT	AAE22226	ck: 7634 len: 230	! Aae22226 Human claudin-D2 protein. 7/20
1	ABG96301	ck: 6154 len: 209	! Abg96301 Human ovarian cancer marker M360.	30: MLLPS	W(K,R)XX(S,A)(Y,F)XG W(K)xx(S)(Y)XG	ASIVT
30: CALPM		W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG	SNIVT	AAE22228	ck: 63 len: 224	! Aae22228 Human claudin-D17 protein. 7/2
1	ABP67991	ck: 8102 len: 230	! Abp67991 Human colon cancer related polypep	30: TLLPQ	W(K,R)XX(S,A)(Y,F)XG W(R)xx(A)(F)XG	SNIVT
30: MLLPS		W(K,R)XX(S,A)(Y,F)XG W(K)xx(S)(Y)XG	ASIVT	ABB95518	ck: 8102 len: 230	! Abb95518 Human angiogenesis related pro
1	ABP65274	ck: 595 len: 569	! Abp65274 Bifidobacterium longum NCC2705 ORF	30: MLLPS	W(K,R)XX(S,A)(Y,F)XG W(K)xx(S)(Y)XG	ASIVT
334: LMIFP		W(R)xx(A)(F)XG WRWFARAG	LPSPS	ABB95541	ck: 4773 len: 220	! Abb95541 Human angiogenesis related pro
1	ABB81045	ck: 2784 len: 8	! Abb81045 Fab fragment directed against clau	30: CALPM	W(K,R)XX(S,A)(Y,F)XG W(K)xx(A)(F)XG	NSIVT
1:		WKIYSYAG		AAO20543	ck: 3152 len: 211	! Aao20543 Human Claudin-1 protein. 6/200
1	AAO15425	ck: 1066 len: 205	! Aao15425 Human gusset metabolic gene (GMG-1	30: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(R)xx(S)(Y)XG	DNIVT
192: NLLGG		WKYSFSG	FLIFP	AAO20545	ck: 4066 len: 211	! Aao20545 Human Claudin-19 protein. 6/20
1	ABG64505	ck: 9468 len: 140	! Abg64505 Human albumin fusion protein #1180	30: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(K)xx(S)(Y)XG	DAIIT
30: MLLPS		W(K,R)XX(S,A)(Y,F)XG W(K)xx(S)(Y)XG	ASIVT	AAO20546	ck: 9812 len: 193	! Aao20546 Murine Claudin-19 protein (par
1	ABG64507	ck: 8102 len: 230	! Abg64507 Human albumin fusion protein #1182	20: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(K)xx(S)(Y)XG	DAIIT
30: MLLPS		W(K,R)XX(S,A)(Y,F)XG W(K)xx(S)(Y)XG	ASIVT	ABP04007	ck: 6970 len: 85	! Abp04007 Human OREF protein sequence SE
1	ABP41557	ck: 5827 len: 401	! Abp41557 Human ovarian antigen HVCAC71, SEQ	57: RDAWV	W(K,R)XX(S,A)(Y,F)XG W(K)xx(S)(F)XG	GKEYM
341: VTGRT		W(K,R)XX(S,A)(Y,F)XG W(K)xx(A)(F)XG	WKSVE	AAU76534	ck: 8102 len: 230	! Aau76534 Tumour-associated antigenic ta

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG ASIVT
30: MLLPS WKTSYVG

ABB90240 ck: 9468 len: 140 ! Abb90240 Human polypeptide SEQ ID NO 2616.
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG ASIVT
30: MLLPS WKTSYVG

ABB91999 ck: 1756 len: 303 ! Abb91999 Herbicidally active polypeptide SE
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(Y)XG MQIGY
15: IPGPK WKVATGG

ABB93775 ck: 275 len: 895 ! Abb93775 Herbicidally active polypeptide SE
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG LPHT
621: RSQGQ WKVVSFAG

ABB84912 ck: 8102 len: 230 ! Abb84912 Human PRO1356 protein sequence SEQ
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG ASIVT
30: MLLPS WKTSYVG

ABB84935 ck: 4773 len: 220 ! Abb84935 Human PRO1488 protein sequence SEQ
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG NSIIV
30: CALPM WKVTAFIG

AAU76231 ck: 2066 len: 211 ! Aau76231 Human senescence associated epithe
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG DNIIV
30: TALPQ WKIVSYAG

AAU83691 ck: 4773 len: 220 ! Aau83691 Human PRO protein, Seq ID No 200.
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG NSIIV
30: CALPM WKVTAFIG

ABE04707 ck: 8102 len: 230 ! Abb04707 Human SP82 protein SEQ ID NO: 2. 3/
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG ASIVT
30: MLLPS WKTSYVG

AAU10521 ck: 8182 len: 230 ! Aau10521 Human CASB81 polypeptide. 2/2002
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG ASIVT
30: MLLPS WKTSYVG

AAU10522 ck: 6271 len: 230 ! Aau10522 Murine CASB81 polypeptide. 2/2002
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG ASIVT
30: MLLPN WKTSYVG

AAU10875 ck: 9502 len: 1,173 ! Aau10875 Penicillium citrinum HMG-CoA re
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG DADVL
97: GPQNG WKWQSFEG

ABG51536 ck: 709 len: 5,701 ! Abg51536 Human liver peptide, SEQ ID No
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG GSKIT
3,952: NVTLK WKPAYDG

ABB56595 ck: 8381 len: 14 ! Abb56595 Human SNP related amino acid se
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG A
6: MLLPS WRTSSYVG

ABE58144 ck: 7060 len: 7,107 ! ABE58144 Drosophila melanogaster polypep
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG GLKVS
5,614: SCMLN WRPPSYDG

ABB62511 ck: 1032 len: 379 ! Abb62511 Drosophila melanogaster polypep
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WRSVS
319: VVGRV WKGSFAG

ABB70767 ck: 1797 len: 619 ! Abb70767 Drosophila melanogaster polypep
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG GIDNS
303: VIAT WKVVSFAG

AAU43488 ck: 1666 len: 50 ! Aau43488 Propionibacterium acnes immunog
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG SSTTR
E: FGSSP WRPWAFRG

ABG00160 ck: 8660 len: 60 ! Abg00160 Novel human diagnostic protein
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG KKKKI
11: CYGQI WKRISFHG

ABG00890 ck: 4548 len: 414 ! Abg00890 Novel human diagnostic protein
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG YPDLS
275: ARNTS WRPNAFWG

ABG04118 ck: 4411 len: 437 ! Abg04118 Novel human diagnostic protein
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG SGTQR
225: SQRPC WKAXAYPG

ABG07792 ck: 7746 len: 731 ! Abg07792 Novel human diagnostic protein

1	407: SINPG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKAVAFIG	AAVPG	1	310: LPGDN	WRVTSYSG	LQORV
1	ABG11899	ck: 2706 len: 243 ! Abg11899 Novel human diagnostic protein #11		1	ABB36684	ck: 709 len: 5,701 ! Abb36684 Peptide #4190 encoded by human	
1	170: LPGDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRVTSYSG	LQORG	3,952: NVTLK	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKCPAYDG	GSKIT	
1	ABG11900	ck: 2958 len: 292 ! Abg11900 Novel human diagnostic protein #11		220: CALPM	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVTAFIG	SNIVT	
1	133: LPGDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRVTSYSG	LQORG	ABB50279	ck: 1473 len: 220 ! Abb50279 Claudin 3 ovarian tumour marke		
1	ABG16462	ck: 7378 len: 473 ! Abg16462 Novel human diagnostic protein #16		29: CALPM	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVSAFAG	SNIT	
1	368: QPLQS	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKVRAFYG	STSQE	ABB50396	ck: 6151 len: 202 ! Abb50396 Human secreted protein encoded		
1	ABG20671	ck: 3301 len: 570 ! Abg20671 Novel human diagnostic protein #20		22: CALPM	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVTAFIG	SNIVT	
1	4: MAL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRGSAYAG	FLALA	ABB50515	ck: 4456 len: 206 ! Abb50515 Human secreted protein encoded		
1	ABG21490	ck: 2706 len: 639 ! Abg21490 Novel human diagnostic protein #21		192: NLLGG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKYSFSG	FLIFP	
1	230: RPGDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRVTSYSG	LQORG	ABB52752	ck: 9451 len: 665 ! Abb52752 Escherichia coli polypeptide 5		
1	ABG22207	ck: 8207 len: 1,130 ! Abg22207 Novel human diagnostic protein #22		148: FWRVL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRFRAYTG	QVIVA	
1	1,017: AFRCI	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKRESYDG	RWETK	ABB52903	ck: 2461 len: 361 ! Abb52903 Escherichia coli polypeptide 5		
1	ABG27980	ck: 9013 len: 254 ! Abg27980 Novel human diagnostic protein #27		35: DEIKY	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRVTSYKG	GEFPE	
1	73: TALPQ	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKQSSYAG	DAIIT	ABB17568	ck: 1476 len: 67 ! Abb17568 Human nervous system related f		
1	ABG28241	ck: 8784 len: 1,331 ! Abg28241 Novel human diagnostic protein #28		47: SWIFR	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKVMAFTG	LMYST	
1	1,077: LFLEH	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRKWAFRG	PGLPR	ABB22021	ck: 709 len: 5,701 ! Abb22021 Protein #4020 encoded by probe		
1	ABG28410	ck: 9642 len: 5,488 ! Abg28410 Novel human diagnostic protein #28		3,952: NVTLK	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKCPAYDG	GSKIT	
1	2,845: DRFPD	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKAGSYLG	QAVKG	AAU09178	ck: 8102 len: 230 ! Aau09178 Human PRO1356 polypeptide. 1/2		
1	ABG29205	ck: 7810 len: 1,286 ! Abg29205 Novel human diagnostic protein #29		30: MLLPS	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKTSSYVG	ASIVT	

1	AAG90500	ck: 6981	len: 301	! Aag90500 C glutamicum protein fragment SEQ	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	EMGLC	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	NEVSD	365:
1	254: AKRW	WKLISFTG	IGSV							
1	AAG91161	ck: 9818	len: 530	! Aag91161 C glutamicum protein fragment SEQ	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	WKLISFTG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG	WREVSFFG	13: RDSGR
1	132: LRGE	WKRSPFG	VEIFG							
1	AAB82494	ck: 2905	len: 376	! Aab82494 Rhodobacter sphaeroides formaldehy	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKRSPFG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	WKTSSYVG	30: MLLPS
1	316: VTGRV	WKGAFGG	ARGRT							
1	AAG75467	ck: 8565	len: 210	! Aag75467 Human colon cancer antigen protein	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG	WRYTAFIG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKVTAFIG	30: CALPM
1	31: CALPM	WRYTAFIG	SNIVT							
1	AAG81335	ck: 587	len: 496	! Aag81335 Human AFP protein sequence SEQ ID	1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRTAFAG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	WKTSSYVG	30: MLLPS
1	4: MAL	WRTAFAG	FLALA							
1	AAE04207	ck: 8102	len: 230	! Aae04207 Human gene 10 encoded secreted pro	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	WRTAFAG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	WKTSSYVG	30: MLLPS
1	30: MLLPS	WRTAFAG	ASIVT							
1	AAE04228	ck: 9468	len: 140	! Aae04228 Human gene 10 encoded secreted pro	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	WRTAFAG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	WKTSSYVG	30: MLLPS
1	30: MLLPS	WRTAFAG	ASIVT							
1	AAE04281	ck: 5423	len: 113	! Aae04281 Human gene 10 encoded secreted pro	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	WRTAFAG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	WKTSSYVG	30: MLLPS
1	3: PS	WRTAFAG	ASIVT							
1	AAE04286	ck: 432	len: 56	! Aae04286 Human gene 10 encoded secreted pro	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	WRTAFAG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	WKTSSYVG	30: MLLPS
1	6: MLLPS	WRTAFAG	ASIVT							
1	AAB92516	ck: 6755	len: 643	! Aab92516 Human protein sequence SEQ ID NO:1	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKLCAPEG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKVTAFIG	30: CALPM
1	632: TLEKR	WKLCAPEG	IKTT							
1	AAB95213	ck: 2647	len: 434	! Aab95213 Human protein sequence SEQ ID NO:1	1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKLCAPEG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKVTAFIG	30: CALPL
1	30: CALPL	WKLCAPEG	NSIVV							

1 ABR47946 ck: 8102 len: 230 ! ABR47946 Human secreted protein, SEQ ID 837
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
WKTSSYVG
30: MLLPS ASIVT

1 ABR48131 ck: 9468 len: 140 ! ABR48131 Human secreted protein, SEQ ID 102
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
WKTSSYVG
30: MLLPS ASIVT

1 ABG74786 ck: 2447 len: 31,267 ! ABG74786 Human RGS11 protein. 6/2003
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
WKRPAVDG
22,972: NVTLK GSKIT

1 ABU71162 ck: 2066 len: 211 ! ABU71162 Human PRO944 protein. 6/2003
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
WRIYSYAG
30: TALPO DNIVT

1 ABU71278 ck: 828 len: 205 ! ABU71278 Human PRO1486 protein. 6/2003
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
WKTSSYVG
192: NLLGG FLIFP

1 ABU71434 ck: 4773 len: 220 ! ABU71434 Human neoplasia inhibiting PRO pol
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
WKTARIG
30: CALPM NSIVV

1 ABU71545 ck: 8102 len: 230 ! ABU71545 Human secreted polypeptide PRO1356
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
WKTSSYVG
30: MLLPS ASIVT

1 ABU71991 ck: 8102 len: 230 ! ABU71991 Novel human secreted and transmembran
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
WKTSSYVG
30: MLLPS ASIVT

1 ABU72148 ck: 8102 len: 230 ! ABU72148 Human PRO polypeptide #40. 6/2003
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
WKTSSYVG
30: MLLPS ASIVT

1 ABU65619 ck: 2066 len: 211 ! ABU65619 Human secreted/transmembrane prote
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
WRIYSYAG
30: TALPO DNIVT

1 ABU65735 ck: 828 len: 205 ! ABU65735 Human secreted/transmembrane prote

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
WKYSSFSG
192: NLLGG FLIFP

1 ABU65952 ck: 2066 len: 211 ! ABU65952 Novel human secreted and transm
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
WRIYSYAG
30: TALPO DNIVT

1 ABU66068 ck: 828 len: 205 ! ABU66068 Novel human secreted and transm
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
WKYSSFSG
192: NLLGG FLIFP

1 ABU66815 ck: 8102 len: 230 ! ABU66815 Human PRO polypeptide #246. 5/2
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
WKTSSYVG
30: MLLPS ASIVT

1 ABU67091 ck: 8102 len: 230 ! ABU67091 Human secreted/transmembrane, p
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
WKTSSYVG
30: MLLPS ASIVT

1 ABU67456 ck: 2066 len: 211 ! ABU67456 Human secreted/transmembrane pr
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
WRIYSYAG
30: TALPO DNIVT

1 ABU67572 ck: 828 len: 205 ! ABU67572 Human secreted/transmembrane pr
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
WKYSSFSG
192: NLLGG FLIFP

1 ABU59896 ck: 8102 len: 230 ! ABU59896 Novel secreted and transmembran
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
WKTSSYVG
30: MLLPS ASIVT

1 ABU61112 ck: 2066 len: 211 ! ABU61112 Human PRO944 polypeptide. 5/200
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
WRIYSYAG
30: TALPO DNIVT

1 ABU65314 ck: 2066 len: 211 ! ABU65314 Human PRO polypeptide #51. 5/20
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
WRIYSYAG
30: TALPO DNIVT

1 ABU65430 ck: 828 len: 205 ! ABU65430 Human PRO polypeptide #167. 5/2
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG

1	192: NLLGG	WKYSFSG	FLIPP	1	ABU57097	ck: 828	len: 205	! Abu57097 Human PRO polypeptide #167. 4/
	ABU37059	ck: 7568	len: 206	! ABj37059 Human breast cancer / ovarian can		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG	WKYSFSG	FLIPP
	27: CALPM	NRVTRFAG	SNIVT		ABR00178	ck: 8102	len: 230	! ABR00178 Human gene 168 encoded secret
	ABU58227	ck: 9728	len: 426	! Abu58227 Soybean stress response protein #2		W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	WKTSSYVG	ASIVT
	346: KDALI	WKIRSFG	GKEYM		ABR00294	ck: 9468	len: 140	! ABR00294 Human gene 168 encoded secret
	ABU58229	ck: 1234	len: 365	! Abu58229 Wheat stress response protein #19		W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	WKTSSYVG	ASIVT
	284: RDMV	WAKSFPG	GKEYM		ABU54408	ck: 695	len: 2,910	! Abu54408 Schizochytrium Orfa protein se
	ABU58450	ck: 2066	len: 211	! Abu58450 Human PRO polypeptide #51. 4/2003		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	WRLDSFG	FLGNV
	30: TALPQ	WRIYSAG	DNIVT		ABU54411	ck: 9992	len: 200	! Abu54411 Schizochytrium ORFA-KS domain
	ABU58566	ck: 828	len: 205	! Abu58566 Human PRO polypeptide #167. 4/2003		W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	WRLDSFG	FLGNV
	192: NLLGG	WKYSFSG	FLIPP		AAO26545	ck: 3601	len: 434	! AAO26545 Human CMP-SA synthetase protei
	ABU55986	ck: 2066	len: 211	! Abu55986 Human secreted/transmembrane prote		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	WKEVAILG	NEVSD
	30: TALPQ	WRIYSAG	DNIVT		ABP78799	ck: 5616	len: 134	! ABp78799 N. gonorrhoeae amino acid sequ
	ABU56102	ck: 828	len: 205	! Abu56102 Human secreted/transmembrane prote		W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	WKGSAFAG	VKGRS
	192: NLLGG	WKYSFSG	FLIPP		ABP80560	ck: 7056	len: 98	! ABp80560 N. gonorrhoeae amino acid sequ
	ABU56524	ck: 2066	len: 211	! Abu56524 Lung cancer-associated polypeptide		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG	WRLSFG	ADEYD
	30: TALPQ	WRIYSAG	DNIVT		ABU10560	ck: 2066	len: 211	! ABu10560 Human secreted/transmembrane p
	ABU56697	ck: 2066	len: 211	! Abu56697 Lung cancer-associated polypeptide		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	WRIYSAG	DNIVT
	30: TALPQ	WRIYSAG	DNIVT		ABU10676	ck: 828	len: 205	! ABu10676 Human secreted/transmembrane p
	ABU56981	ck: 2066	len: 211	! Abu56981 Human PRO polypeptide #51. 4/2003		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	WKYSFSG	FLIPP
	30: TALPQ	WRIYSAG	DNIVT		ABU05327	ck: 7694	len: 197	! Abu05327 Human diagnostics and therapeu
						W(K,R)XX(S,A)(Y,F)XG		

107: HLFY W(R)xx(A)(Y)XG YEEAI
WEMLAYK

Databases searched:
Geneseq-AA, Release 13.0, Released on 19Jun2003, Formatted on 15Jul2003
Total finds: 271
Total length: 158,726,570
Total sequences: 1,107,863
CPU time: 05:15.75

```
! FINDPATTERNS on genpept:* allowing 0 mismatches
!      1 W(K,R)XX(S,A)(Y,F)XG      August 15, 2003 10:56 ..
CAA23453 ck: 3857 len: 159 ! GB:V00140 reading frame II [Cauliflower mosaic virus]
61: SLLGI WKINSYFG LSKDP
1
CAA23457 ck: 3030 len: 159 ! GB:V00141 reading frame (II) [Cauliflower mosaic virus]
61: SLLGI WKINSYFG LSKDP
1
CAA21733 ck: 3141 len: 159 ! GB:M90543 aphid transmission protein [Cauliflower mosaic virus]
61: SLLGI WKINSYFG LSKDP
1
AAA46346 ck: 3377 len: 159 ! GB:M10376 ORF2; putative [Cauliflower mosaic virus]
61: SLLGI WKINSYFG LSKDP
1
AAA46355 ck: 3038 len: 159 ! GB:M90541 aphid transmission protein [Cauliflower mosaic virus]
61: SLLGI WKINSYFG LSKDP
1
AAA46364 ck: 8557 len: 99 ! GB:M94887 ORF II [Cauliflower mosaic virus]
61: SLLGI WKINSYFG LSKDP
1
AAA60883 ck: 405 len: 146 ! GB:L22579 homolog of vaccinia virus CDS A28
73: DVKQK WRCVAYFG NGFVS
1
AAA62372 ck: 3139 len: 159 ! GB:M90542 aphid transmission protein [Cauliflower mosaic virus]
61: SLLGI WKINSYFG LSKDP
1
AAA66604 ck: 8680 len: 99 ! GB:M17415 unknown protein [Cauliflower mosaic virus]
61: SLLGI WKINSYFG LSKDP
1
AAA96696 ck: 3686 len: 159 ! GB:M37582 ORF II [Cauliflower mosaic virus]
61: SLLGI WKINSYFG LSKDP
1
AAA96697 ck: 3023 len: 159 ! GB:M37581 ORF II [Cauliflower mosaic virus]
61: SLLGI WKINSYFG LSKDP
1
BAB83514 ck: 8144 len: 547 ! GB:AB074892 capsid protein [Swine calicivirus]
410: QDFNQ WRLPAYFG ALTNN
1
AAA46340 ck: 8743 len: 64 ! GB:M32809 aphid acquisition factor [Cauliflower mosaic virus]
7: SLLGI WKINSYFG LSKDP
1
CAA47502 ck: 405 len: 146 ! GB:X67115 A28L [Variola virus] (ver 1)
73: DVKQK WRCVAYFG NGFVS
1
CAA53857 ck: 405 len: 146 ! GB:X76266 ORF5L [Variola virus] (ver 1)
73: DVKQK WRCVAYFG NGFVS
1
CAA49076 ck: 405 len: 146 ! GB:X69198 A31L [Variola virus] (ver 1)
73: DVKQK WRCVAYFG NGFVS
1
BAA82378 ck: 820 len: 955 ! GB:AP000367 EST C28952(C62945) corresponding protein
115: KEDCP WRVHAYKG KWNDY
1
CAA48161 ck: 583 len: 386 ! GB:X68020 SFA [Saccharomyces cerevisiae]
322: VTGRV WKGAATFG IKGRS
1
AAA34344 ck: 8341 len: 381 ! GB:M58332 encoding formaldehyde resistant protein
320: VTGRT WKGAATFG VKGRS
1
BAB92078 ck: 7458 len: 602 ! GB:AB074427 ferredoxin-nitrite reductase
394: LIDPS WKRSTFG VNPQK
1
BAA32244 ck: 8858 len: 777 ! GB:AB008746 transposase-like protein [
```

1	402: ERPEI	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG WRKQSPFG	KLHNF		319: VTGRV	WKGTAFFG	FKSRS	
	BAB09746	ck: 275 len: 895	! GB:AB025621 receptor protein kinase-like [A		CAA98742	ck: 583 len: 386	! GB:Z74216 ORF YDL168w [Saccharomyces cer	
1	621: RSQGG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG WKMVSFAG	LPHFT		322: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGSAFFG	IKGRS	
	BAA31937	ck: 7849 len: 934	! GB:AB012603 HMG-CoA reductase [Pichia jadin		CAB87206	ck: 5672 len: 817	! GB:AL163640 ubiquitin fusion degradation	
1	79: EDGTV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRSRAYHG	KLKGY		706: VFRTI	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WKQPAYFG	KLEQE	
	BAB09054	ck: 5269 len: 379	! GB:AB006703 alcohol dehydrogenase (EC 1.1.1		CAA21785	ck: 1472 len: 380	! GB:AL032681 probable glutathione-depende	
1	317: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKGTAFFG	FKSRT		319: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGCAFFG	VKGRS	
	BAA10929	ck: 2338 len: 530	! GB:D64052 cytochrome P450 like_TBP [Nicotia		CAA97989	ck: 8779 len: 475	! GB:Z73615 ORF YPL259c [Saccharomyces cer	
1	254: RGCSF	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRPDATFG	GNDYY		382: KSAIL	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKIRSFFG	GKEYS	
	AAB06322	ck: 5323 len: 379	! GB:U63931 glutathione-dependent formaldehyd		AAC35913	ck: 5085 len: 379	! GB:AF066054 formaldehyde dehydrogenase [
1	317: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKGTAFFG	FKSRT		318: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGCAFFG	IKGRT	
	CAC19740	ck: 3117 len: 1,010	! GB:AL512487 ubiquitin fusion degradation pr		BAC15531	ck: 8356 len: 1,727	! GB:AP003849 similar to mutator-like tran	
1	706: VFRTI	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WKQPAYFG	KLBOE		266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	
	CAA91578	ck: 583 len: 386	! GB:Z67750 alcohol dehydrogenase [Saccharomy		BAC16016	ck: 3759 len: 903	! GB:AP004308 contains ESTs C28952(C62945)	
1	322: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKGSAFFG	IKGRS		266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	
	CAD21500	ck: 9212 len: 380	! GB:AL670542 probable alcohol dehydrogenase		BAC16392	ck: 5613 len: 1,179	! GB:AP003749 putative mutator-like transp	
1	319: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGSAFFG	VKGRS		273: KDCCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWKDY	
	CAB42828	ck: 7851 len: 474	! GB:X60288 medium chains of clathrin associa		AAD51707	ck: 8565 len: 274	! GB:AF174004 alternative oxidase [Triticu	
1	382: KSAIL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG WKIRSFFG	GKEYS		21: GRRRR	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRISSTWG	IEQSK	
	AAB19117	ck: 4843 len: 381	! GB:U77637 class III ADH enzyme [Oryza sativ		BAC20901	ck: 8873 len: 1,753	! GB:AP005764 contains ESTs C28952(C62945)	
1		W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG			266: KEDCP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	KWNDY	

1 BAC21493 ck: 3759 len: 903 ! GB:AP004988 contains ESTs C28952(C62945),AU
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
266: KEECP WRVHAYKG KWNDY

1 BAC21600 ck: 9400 len: 1,662 ! GB:AP005516 contains ESTs C28952(C62945),AU
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
266: KEECP WRVHAYKG KWNDY

1 BAC20567 ck: 9502 len: 1,173 ! GB:AB072893 HMG-CoA reductase [Penicillium
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
97: GPONG WKWSFDG DADVL

1 BAA92398 ck: 5721 len: 1,281 ! GB:AP001366 unnamed protein product [Oryza
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
237: KEECP WRVHAYKG KWNDY

1 BAA92402 ck: 51 len: 938 ! GB:AP001366 unnamed protein product [Oryza
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
266: KEECP WRVHAYKG KWNDY

1 BAB03445 ck: 5721 len: 1,281 ! GB:AP002817 unnamed protein product [Oryza
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
237: KEECP WRVHAYKG KWNDY

1 BAB03449 ck: 51 len: 938 ! GB:AP002817 unnamed protein product [Oryza
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
266: KEECP WRVHAYKG KWNDY

1 BAC04691 ck: 9234 len: 224 ! GB:AK096063 unnamed protein product [Homo s
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
30: TALPQ WKQSIAG DAIIT

1 AAA51596 ck: 9148 len: 374 ! GB:M81118 alcohol dehydrogenase [Homo sapie
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
314: VTGRT WKGTAFGG WKSVE

1 AAA51597 ck: 9878 len: 392 ! GB:M29872 alcohol dehydrogenase class III I
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
332: VTGRT WKGTAFGG WKSVE

1 BAB21814 ck: 920 len: 1,554 ! GB:AB051510 KIAA1723 protein [Homo sapiens]
W(K,R)XX(S,A)(Y,F)XG

1 864: NGSVN W(R)XX(S)(F)XG
WRTGSFHG PGHIS

1 BAA22984 ck: 6154 len: 209 ! GB:AB000712 CPE-receptor [Homo sapiens]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
30: CALPM WRVTAFIG SNIVT

1 BAA22986 ck: 1473 len: 220 ! GB:AB000714 RVP1 [Homo sapiens] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
29: CALPM WRVSFIG SNIVT

1 AAA79018 ck: 9148 len: 374 ! GB:M30471 alcohol dehydrogenase 3 [Homo
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
314: VTGRT WKGTAFGG WKSVE

1 BAA22781 ck: 6200 len: 209 ! GB:D88492 CPE-receptor [Cercopithecus a
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
30: CALPM WRVTAFIG SNIVT

1 BAB62195 ck: 1821 len: 414 ! GB:AB066508 hypothetical protein [Macac
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
403: TLEKR WKLCFEG IXTT

1 AAB48880 ck: 8779 len: 75 ! GB:U61018 T-cell receptor beta [Sagulinu
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
57: STLRG WKGEAFPG EGTKL

1 AAB81637 ck: 5589 len: 1,083 ! GB:AF026219 HP protein [Homo sapiens] (
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
393: NGSVN WRTGSFHG PGHIS

1 AAB87700 ck: 9802 len: 1,091 ! GB:AF035119 deleted in liver cancer-1 [
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
401: NGSVN WRTGSFHG PGHIS

1 BAB47150 ck: 4720 len: 432 ! GB:AB027414 CMP-sialic acid synthetase
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
355: DKOLD WKEVAYMG NDAPD

1 AAB41888 ck: 6805 len: 376 ! GB:U84791 alcohol dehydrogenase class I
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
316: VTGRV WKGTAFGG WKSVE

901: LFGDN WRVTSYSG LQORG
 1
 BAA15667 ck: 7831 len: 261 ! GB:D90828 ORF_ID:o336gap#12-similar to [Swi
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 28: GSEVV WRRMSYFG DTLAH
 1
 BAA15670 ck: 7831 len: 261 ! GB:D90829 ORF_ID:o336gap#12-similar to [Swi
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 28: GSEVV WRRMSYFG DTLAH
 1
 BAB505321 ck: 2788 len: 540 ! GB:AF001512 D-3-phosphoglycerate dehydrogen
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 141: IKAGE WKRFQFG TELRG
 1
 BAA22412 ck: 7661 len: 369 ! GB:D38504 formaldehyde dehydrogenase [Esche
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 309: VTGRV WKSFAFG VKGRS
 1
 BAB52832 ck: 8644 len: 700 ! GB:AF003009 unknown protein [Mesorhizobium
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 379: ELLGP WKRFQFG AGAGT
 1
 BAB53265 ck: 5716 len: 412 ! GB:AF003011 sugar ABC transporter, periplas
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(F)XG
 315: NIEAA WRVSWFG QTDAA
 1
 BAB54248 ck: 9706 len: 219 ! GB:AF003012 unknown protein [Mesorhizobium
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 148: CMDAA WRKSING KOORI
 1
 BAB53928 ck: 2074 len: 344 ! GB:AF003013 similar to O-acyltransferase [M
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 264: ROIKI WRSVYIG LISYP
 1
 AAA65962 ck: 9499 len: 375 ! GB:L36327 glutathione-dependent formaldehyd
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 315: VTGRV WKGFQFG ARGRT
 1
 AAB81031 ck: 2467 len: 92 ! GB:U38702 YebI [Escherichia coli] (ver 1)
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 28: GSPVV WRRMSYFG DTLAH
 1

AAB06701 ck: 5809 len: 213 ! GB:U33003 25 kDa outer membrane protein
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 74: IKPDD WKAGAFAG WNFQQ
 1
 AAB06702 ck: 2394 len: 201 ! GB:U33004 25 kDa outer membrane protein
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 74: IKPDD WKAGAFAG WNFQQ
 1
 AAB09774 ck: 2905 len: 376 ! GB:L47326 alcohol dehydrogenase [Rhodoba
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 316: VTGRV WKSFAFG ARGRT
 1
 AAB36692 ck: 6016 len: 213 ! GB:U39358 25 kDa outer membrane protein
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 74: IKPDD WKAGAFAG WNFQQ
 1
 AAB36694 ck: 5851 len: 213 ! GB:U39359 25 kDa outer membrane protein
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 74: IKPDD WKAGAFAG WNFQQ
 1
 AAB36695 ck: 5911 len: 213 ! GB:U39397 25 kDa outer membrane protein
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 74: IKPDD WKAGAFAG WNFQQ
 1
 AAB40467 ck: 1854 len: 1,180 ! GB:U29591 exonuclease V subunit [Escheri
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 901: LFGDN WRVTSYSG LQORG
 1
 AAB61921 ck: 224 len: 369 ! GB:U89767 glutathione dependent formalde
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 309: VTGRV WKSFAFG ARGRT
 1
 AAB99205 ck: 6435 len: 70 ! GB:U67562 LSU ribosomal protein L24E [Me
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(F)XG
 4: MPE WFTCSFG YEIEP
 1
 AAC21854 ck: 2043 len: 378 ! GB:U32703 alcohol dehydrogenase, class I
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 318: VTGRV WKSFAFG VKGRS
 1
 AAC22066 ck: 1426 len: 261 ! GB:U32724 conserved hypothetical integra
 W(K,R)XX(S,A)(Y,F)XG

28: GVEVV	W(R)xx(A)(Y)XG WKMAYFG	DTLSH	1	AAF12100	ck: 3668	len: 206	! GB:AE002085 3-demethylubiquinone-9 3-me-
AAC44551	ck: 9499	len: 375	! GB:U34346 NAD-GSH-dependent formaldehyde de	111: VEETG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG	WRTASFDG	VVWGA
315: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	ARGRT	1	AAD56369	ck: 1854	len: 1,180	! GB:AF179304 RecB2109 [Escherichia coli]
BAB98009	ck: 387	len: 275	! GB:AP005275 Hypothetical protein [Corynebac	901: LFGDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	WERTSYSG	LQORG
130: AAVIG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	WIDSG	1	BAC17437	ck: 419	len: 254	! GB:AP005216 hypothetical protein [Coryn-
BAB98072	ck: 6981	len: 301	! GB:AP005276 Hypothetical membrane protein [114: AAMIG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WREFAYAG	VVESG
254: AKRNW	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	IGSVV	1	BAC18189	ck: 8770	len: 530	! GB:AP005218 putative D-3-phosphoglycera
BAB98677	ck: 9818	len: 530	! GB:AP005278 Phosphoglycerate dehydrogenase	132: LRDSG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	WRSFNG	VEIFG
132: LREGE	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	VEIFG	1	BAA14135	ck: 2790	len: 513	! GB:D90108 aqualysin precursor [Thermus
BAC08744	ck: 1651	len: 399	! GB:AP005373 oxygen independent coprophorphy	495: AGYYL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRTAISG	SGMYE
258: HNQVY	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	FGMGA	1	BAA12834	ck: 3701	len: 366	! GB:D85613 30KD protein, similar to ADHX
AAD35451	ck: 982	len: 441	! GB:AE001716 4-alpha-glucanotransferase [The	306: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKGSAFGG	VKGRS
348: EGQTF	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	PFSGI	1	BAB93573	ck: 6278	len: 369	! GB:AP004237 alcohol dehydrogenase [Prot.
BAA29818	ck: 9251	len: 278	! GB:AP000003 278aa long hypothetical protein	309: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKGSAFGG	VKGRT
99: PERVG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	ISFQN	1	BAC59114	ck: 8487	len: 261	! GB:AP005075 zinc ABC transporter, perme
AAB18081	ck: 3701	len: 366	! GB:U73857 30 kD protein, similar to ADHX-RA	28: GSFVV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WKMAYFG	DTLAH
306: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	VKGRS	1	BAC60637	ck: 3903	len: 1,227	! GB:AP005081 exodeoxyribonuclease V, 135
AAD46808	ck: 1674	len: 1,083	! GB:AF157643 RecB [Mycobacterium smegmatis]	930: PIDRN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	WRTSYSG	LVKQG
766: AIDMA	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG	LLRAA	1	BAC61414	ck: 3100	len: 382	! GB:AP005084 putative alcohol dehydrogen
BAC24261	ck: 9466	len: 265	! GB:AB063521 yebI [Wigglesworthia brevipalpi]	320: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG	WKGSAFGG	VKGRS
28: GSFLI	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	DTLSH	1	BAC61619	ck: 9481	len: 904	! GB:AP005084 putative integral membrane

1 1 W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG LGASL
872: SLDFG WRAISFLG
BAC45296 ck: 4392 len: 125 ! GB:AF005935 bil10031 [Bradyrhizobium japonicum]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG PFICG
100: SCDDA WRVAFNG
BAC45788 ck: 548 len: 275 ! GB:AF005936 bil10523 [Bradyrhizobium japonicum]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG AVVFG
185: LAPPD WRPLAYAG
AAA20658 ck: 7421 len: 663 ! GB:U04331 leukocyte-type 12-lipoxygenase [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG YQFLN
223: RVVNS WKEDAFNG
AAA41532 ck: 6102 len: 663 ! GB:L06040 12-lipoxygenase [Rattus norvegicus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG YQFLN
223: RVVNS WKEDAFNG
AAA41760 ck: 2015 len: 280 ! GB:M74067 CDS is longest ORF of a cDNA clone
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG SSIIT
29: CALPM WRVSFAFNG
AAA64930 ck: 7262 len: 663 ! GB:L34570 12-lipoxygenase [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG YQFLN
223: RVVNS WKEDAFNG
AAA68896 ck: 1156 len: 374 ! GB:M84147 alcohol dehydrogenase-B2 [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG WKSVE
314: VTGRT WKGTAFFG
BAA22985 ck: 1978 len: 210 ! GB:AB000713 CPE-receptor [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG SNIIV
30: CALPM WRVTAFIG
BAA21675 ck: 6032 len: 1,093 ! GB:D31962 RhoGAP [Rattus rattus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG PGHLS
393: NSSVN WRTGSFNG
AAB30132 ck: 6502 len: 663 ! GB:S63383 12-lipoxygenase [Rattus sp.] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG YQFLN
223: RVVNS WKEDAFNG

AAC27078 ck: 2898 len: 211 ! GB:AF072127 claudin-1 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG DNIIVT
30: TALPQ WKIYSYAG
AAC27079 ck: 5271 len: 230 ! GB:AF072128 claudin-2 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG ASIIVT
30: MLLPN WRTSSYVG
AAC52763 ck: 1156 len: 374 ! GB:U48970 class III alcohol dehydrogenase
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG WKSVE
314: VTGRT WKGTAFFG
AAD09756 ck: 7312 len: 219 ! GB:AF087821 claudin-3 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG SSIIT
29: CALPM WRVSFAFNG
AAD09757 ck: 1978 len: 210 ! GB:AF087822 claudin-4 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG SNIIVT
30: CALPM WRVTAFIG
AAD09759 ck: 3382 len: 219 ! GB:AF087824 claudin-6 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG NSIIV
30: CALPM WKVTAFIG
AAD14608 ck: 7312 len: 219 ! GB:AF095905 CPETR2 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG SSIIT
29: CALPM WRVSFAFNG
AAD17319 ck: 9358 len: 217 ! GB:AF124424 claudin-9 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG NSIIV
30: CALPL WKVTAFIG
AAD51760 ck: 234 len: 1,092 ! GB:AF178078 deleted in liver cancer 1 [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG PGHLS
402: NSSVN WRTGSFNG
AAP36374 ck: 9118 len: 210 ! GB:BT007706 Homo sapiens claudin 4 [synt]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG SNIIVT
30: CALPM WRVTAFIG
CAD48659 ck: 7634 len: 230 ! GB:AX497200 unnamed protein product [Homo sapiens]

1 BAC30640 ck: 2898 len: 211 ! GB:AK040604 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
30: TALPQ WKIISYAG DNIYT

1 BAC31288 ck: 241 len: 305 ! GB:AK042544 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
198: QKQPA WKVGAYKG KQAIS

1 BAB23725 ck: 6271 len: 230 ! GB:AK004990 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
30: MLLPN WRTSSYVG ASIYT

1 BAC33296 ck: 890 len: 224 ! GB:AK048287 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
30: TLIPQ WRVSAFIG SNIII

1 BAC33845 ck: 675 len: 689 ! GB:AK049628 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
384: QLMEN WKHSHYAG EQPET

1 BAC34616 ck: 3107 len: 490 ! GB:AK051372 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
198: QKQPA WKVGAYKG KQAIS

1 BAC34782 ck: 4917 len: 643 ! GB:AK051821 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
632: TLEKR WKICAFEG IKTT

1 BAC35693 ck: 2898 len: 211 ! GB:AK054207 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
30: TALPQ WKIISYAG DNIYT

1 BAC38267 ck: 2898 len: 211 ! GB:AK081601 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
30: TALPQ WKIISYAG DNIYT

1 BAC39813 ck: 6000 len: 200 ! GB:AK087150 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
131: EMGLC WKEVAYLG NEVSD

1 BAB22275 ck: 7312 len: 219 ! GB:AK002672 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
1 BAC39813 ck: 6000 len: 200 ! GB:AK087150 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
131: EMGLC WKEVAYLG NEVSD

29: CALPM W(R)XX(A)(F)XG SSIIT
WRVSAFIG

1 BAB27028 ck: 3382 len: 219 ! GB:AK010560 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
30: CALPM WKVTAFIG NSIIV

1 BAB27114 ck: 3382 len: 219 ! GB:AK010682 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
30: CALPM WKVTAFIG NSIIV

1 BAB28403 ck: 7312 len: 219 ! GB:AK012677 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
29: CALPM WRVSAFIG SSIIT

1 BAC36370 ck: 221 len: 374 ! GB:AK076507 unnamed protein product [Mus musculus]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
314: VTGRT WKGTAFGG WKSVE

1 BAC43063 ck: 3045 len: 265 ! GB:AK118455 unknown protein [Arabidopsis thaliana]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
210: PQQIN WRPPSYFG RDDLE

1 BAC55711 ck: 111 len: 1,656 ! GB:AP004275 P0453E05.27 [Oryza sativa (L.) indica]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(Y)XG
266: KEECP WRVHAYKG KWDY

1 BAC16635 ck: 4488 len: 380 ! GB:AB085186 formaldehyde dehydrogenase [Homo sapiens]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
318: VTGRV WRGCAFGG VKGRT

1 BAC22123 ck: 8797 len: 230 ! GB:AB060689 putative esterase [Gibberella fujikuroi]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
170: IDTDD WRYSFSG AEPPL

1 BAA95853 ck: 2865 len: 1,591 ! GB:AP002070 unnamed protein product [Oryza sativa (L.) indica]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(Y)XG
266: KEDCP WRVHAYKG KWDY

1 BAA96559 ck: 2865 len: 1,591 ! GB:AP002480 unnamed protein product [Oryza sativa (L.) indica]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(Y)XG
266: KEDCP WRVHAYKG KWDY

1	BAB61368	ck: 3033 len: 1,626 ! GB:AP003315 putative mutator-like transposase	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVHAYKG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVHAYKG	742: AVEDN WRVHAYKG AGDEQ
1	266: KEDCP		KWNDY		
1	BAB44014	ck: 2586 len: 429 ! GB:AP003023 putative clathrin-associated protein	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKVKSFPG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTLSTQG	BAB90249 ck: 5321 len: 542 ! GB:AP003412 B1150F11.9 [Oryza sativa (j 247: FGRC KLFMV
1	347: RDMV		GRDYM		
1	BAB89726	ck: 5321 len: 542 ! GB:AP003269 P0504E02.28 [Oryza sativa (japo	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKVKSFPG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	BAB92621 ck: 4995 len: 1,592 ! GB:AP003416 putative mutator-like trans 266: KEDCP KWNDY
1	247: FGRC		KLFMV		
1	BAB90682	ck: 5760 len: 1,110 ! GB:AP004225 putative mutator-like transposase	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	BAB90513 ck: 1142 len: 1,604 ! GB:AP003760 putative mutator-like trans 266: KEDCP KWNDY
1	209: KEDCP		KWNDY		
1	BAB61148	ck: 1244 len: 603 ! GB:AP003199 B1045D11.16 [Oryza sativa (japo	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKVAFQG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	BAB89480 ck: 4528 len: 1,353 ! GB:AP003220 putative mutator-like trans 265: KEDCP KWNDY
1	410: QTNLS		KIYAA		
1	BAB92344	ck: 1421 len: 1,605 ! GB:AP003273 putative mutator-like transposase	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG WRAVSFG	BAC22307 ck: 2309 len: 268 ! GB:AP003884 OJ1136_AL0.3 [Oryza sativa 243: FFRSF ARRSV
1	266: KEDCP		KWNDY		
1	BAB85342	ck: 4807 len: 929 ! GB:AP003448 putative mutator-like transposase	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	BAC64991 ck: 5936 len: 1,680 ! GB:AP004461 P0443G08.15 [Oryza sativa (266: KEDCP KWNDY
1	BAB84495	ck: 1142 len: 1,604 ! GB:AP004194 putative mutator-like transposase	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	BAC55638 ck: 159 len: 1,618 ! GB:AP003865 OJ1081_B12.20 [Oryza sativa 266: KEDCP KWNDY
1	BAB91969	ck: 5760 len: 1,110 ! GB:AP004258 putative mutator-like transposase	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRVHAYKG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG W(K)XX(A)(F)XG WKGTAFFG	CAA71913 ck: 4477 len: 381 ! GB:Y11029 glutathione-dependent formalde 319: VTGRV FKSRT
1	209: KEDCP		KWNDY		
1	BAB92293	ck: 2586 len: 429 ! GB:AP003245 putative clathrin-associated protein	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKVKSFPG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKVDAFTG	CAA50298 ck: 845 len: 396 ! GB:X70962 acyl-[acyl-carrier protein] d 340: FLVGR LSGEG
1	347: RDMV		GRDYM		
1	BAB89913	ck: 8647 len: 1,080 ! GB:AP003300 putative anthranilate phosphori		W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRPPSYFG	CAB38623 ck: 7947 len: 700 ! GB:AL035556 putative protein [Arabidops 645: PQGIN RDLLE

1
CAA57973 ck: 5257 len: 379 ! GB:X82647 class III ADH, glutathione-depend
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(F)XG
317: VTGRV WKGTAFGG FKST

1
CAB80552 ck: 7947 len: 700 ! GB:AL161594 putative protein [Arabidopsis t
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(Y)XG
645: PQQIN WRPPSYFG RDDL

1
CAD39292 ck: 5056 len: 1,421 ! GB:AL606992 OSJNBa0073L13.8 [Oryza sativa
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(Y)XG
147: KKDCP WRVYAYKG KWDY

1
CAA07349 ck: 9934 len: 396 ! GB:AJ006957 stearyl-acyl carrier protein d
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(F)XG
340: FLVGR WKVDAFTG LSGEG

1
CAA07350 ck: 9787 len: 396 ! GB:AJ006958 stearyl-acyl carrier protein d
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(F)XG
340: FLVGR WKVDAFTG LSGEG

1
BAC25945 ck: 2898 len: 211 ! GB:AK028428 unnamed protein product [Mus mu
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(S)(Y)XG
30: TALPQ WKIYSYAG DNIVT

1
AAD22962 ck: 1690 len: 211 ! GB:AF115546 claudin-1 [Homo sapiens] (ver 1
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(Y)XG
30: TALPQ WKIYSYAG DNIVT

1
CAD23055 ck: 8102 len: 230 ! GB:AL158821 dJ75H8.1 (claudin 2) [Homo sapi
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(S)(Y)XG
30: MLLPS WKTSSYVG ASIVT

1
BAA09078 ck: 8890 len: 1,220 ! GB:D50487 RNA helicase [Homo sapiens] (ver
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(F)XG
538: NDIPE WKHAFGG NKASY

1
CAA10254 ck: 8813 len: 217 ! GB:AJ130941 claudin-9 protein [Homo sapiens
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(F)XG
30: CALPL WKVTAFIG NSIVV

1
CAB60616 ck: 63 len: 224 ! GB:AJ250712 claudin-17 [Homo sapiens] (ver

1
30: TLLPQ W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(F)XG
WRVSAFVG SNIIV

CAD12455 ck: 404 len: 26,926 ! GB:AJ277892 N2B-Titin Isoform [Homo sap
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(Y)XG
18,126: NVTLK WKKPAYDG GSKIT

1
CAD12456 ck: 1298 len: 34,350 ! GB:AJ277892 Titin [Homo sapiens] (ver 1
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(Y)XG
25,550: NVTLK WKKPAYDG GSKIT

1
CAB56533 ck: 5150 len: 220 ! GB:AJ249735 claudin-6 [Homo sapiens] (ve
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(F)XG
30: CALPM WKVTAFIG NSIVV

1
BAA95566 ck: 63 len: 224 ! GB:AP001707 CLDN17 [Homo sapiens] (ver 1
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(F)XG
30: TLLPQ WRVSAFVG SNIIV

1
AAE76203 ck: 2647 len: 434 ! GB:AF271388 CMP-N-acetylneuraminic acid
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(Y)XG
365: EMGLC WKEVAYLG NEVSD

1
CAR62188 ck: 227 len: 26,926 ! GB:X90568 Protein sequence and annotati
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(Y)XG
18,126: NVTLK WKKPAYDG GSKIT

1
AAG17984 ck: 8102 len: 230 ! GB:AF177340 claudin 2 [Homo sapiens] (ve
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(S)(Y)XG
30: MLLPS WKTSSYVG ASIVT

1
AAF61393 ck: 2066 len: 211 ! GB:AF134160 claudin-1 [Homo sapiens] (ve
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(S)(Y)XG
30: TALPQ WFIYSYAG DNIVT

1
AAK02013 ck: 5150 len: 220 ! GB:AF125306 Skullin 2 [Homo sapiens] (ve
W(K,R)XX(S,A)(Y,F)XG
W(K)xx(A)(F)XG
30: CALPM WKVTAFIG NSIVV

1
BAC51480 ck: 9211 len: 369 ! GB:AP005957 alcohol dehydrogenase class
W(K,R)XX(S,A)(Y,F)XG
W(R)xx(A)(F)XG

415: SIETD W(R)XX(A)(F)XG YLSPY
 W(R)XX(A)(F)XG
 CAA34817 ck: 9898 len: 437 ! GB:X16944 48kd protein (AA 1 to 437) [Esche
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 279: SIETD W(R)XX(A)(F)XG YLSPY
 W(R)XX(A)(F)XG
 CAA34818 ck: 7330 len: 303 ! GB:X16944 33kd protein (AA 1 to 303) [Esche
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 145: SIETD W(R)XX(A)(F)XG YLSPY
 W(R)XX(A)(F)XG
 CAA34819 ck: 5402 len: 182 ! GB:X16944 20kd protein (AA 1 to 182) [Esche
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 24: SIETD W(R)XX(A)(F)XG YLSPY
 W(R)XX(A)(F)XG
 CAC46492 ck: 5059 len: 275 ! GB:AL591788 PROBABLE HIGH-AFFINITY ZINC UPT
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 29: GCFVI W(R)XX(A)(F)XG DTMAR
 W(R)XX(A)(F)XG
 CAC31246 ck: 3832 len: 367 ! GB:AL583920 putative aminomethyltransferase
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 263: GWAIG W(K)XX(A)(F)XG RDALL
 W(K)XX(A)(F)XG
 CAC30645 ck: 421 len: 528 ! GB:AL583923 D-3-phosphoglycerate dehydrogen
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(S)(F)XG
 129: LRAHI W(K)XX(S)(F)XG TEIFG
 W(K)XX(S)(F)XG
 CAB16440 ck: 421 len: 528 ! GB:Z99263 phosphoglycerate dehydrogenase [M
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(S)(F)XG
 129: LRAHI W(K)XX(S)(F)XG TEIFG
 W(K)XX(S)(F)XG
 CAB11378 ck: 3832 len: 367 ! GB:Z98741 aminomethyltransferase [Mycobacte
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 263: GWAIG W(K)XX(A)(F)XG RDALL
 W(K)XX(A)(F)XG
 CAA28250 ck: 1854 len: 1,180 ! GB:X04581 exonuclease V (AA 1-1180) [Esche
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 901: LPGDN W(R)XX(S)(Y)XG LQQRG
 W(R)XX(S)(Y)XG
 AAG21801 ck: 9791 len: 431 ! GB:AF192346 aspartate transcarbamoylase [Bu
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(Y)XG
 414: MRDAT W(R)XX(A)(Y)XG PEDAV
 W(R)XX(A)(Y)XG

CAD55617 ck: 4956 len: 369 ! GB:X04616 putative alcohol dehydrogenase
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 309: VTGRV W(R)XX(A)(F)XG VKGRS
 W(R)XX(A)(F)XG
 CAC89809 ck: 8384 len: 765 ! GB:AJ414145 putative kinase [Yersinia pe
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 306: ADAPL W(R)XX(S)(Y)XG GKGVV
 W(R)XX(S)(Y)XG
 CAC89863 ck: 4557 len: 1,220 ! GB:AJ414146 exodeoxyribonuclease V beta
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 904: KMODY W(R)XX(S)(Y)XG LQQRG
 W(R)XX(S)(Y)XG
 CAC90298 ck: 7342 len: 766 ! GB:AJ414148 hypothetical protein [Yersin
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 306: ADAPL W(R)XX(S)(Y)XG GKGVV
 W(R)XX(S)(Y)XG
 CAC90325 ck: 6334 len: 377 ! GB:AJ414148 probable alcohol dehydrogena
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 310: VTGRV W(R)XX(A)(F)XG VKGRS
 W(R)XX(A)(F)XG
 CAC90871 ck: 7586 len: 261 ! GB:AJ414151 high-affinity zinc uptake sy
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 28: GSFVV W(R)XX(S)(Y)XG DTLAH
 W(R)XX(S)(Y)XG
 AAC73459 ck: 7661 len: 369 ! GB:AE000142 alcohol dehydrogenase class
 W(K,R)XX(S,A)(Y,F)XG
 W(K)XX(A)(F)XG
 309: VTGRV W(K)XX(A)(F)XG VKGRS
 W(K)XX(A)(F)XG
 AAC74929 ck: 7831 len: 261 ! GB:AE000280 orf, hypothetical protein [E
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 28: GSFVV W(R)XX(S)(Y)XG DTLAH
 W(R)XX(S)(Y)XG
 AAC75859 ck: 1854 len: 1,180 ! GB:AE000365 DNA helicase, ATP-dependent
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(S)(Y)XG
 901: LPGDN W(R)XX(S)(Y)XG LQQRG
 W(R)XX(S)(Y)XG
 CAB64351 ck: 3052 len: 372 ! GB:AJ243941 putative formaldehyde dehydr
 W(K,R)XX(S,A)(Y,F)XG
 W(R)XX(A)(F)XG
 310: VTGRV W(R)XX(A)(F)XG VKGRT
 W(R)XX(A)(F)XG
 CAB62689 ck: 4085 len: 660 ! GB:AL939107 putative regulator. [Strepto

1	135: AALGE	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRGPAYAG	FGAAD	CAB52011	ck: 1160 len: 330	! GB:AL939111 putative membrane protein [Streptococcus pneumoniae]	1	AAF32314	ck: 2510 len: 197	! GB:AF218379 CBLN3 [Mus musculus] (ver 1)
1	31: KARLR	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WELAFAG	FVGVV	CAB62714	ck: 398 len: 65	! GB:AL939112 hypothetical protein SC4A7.10 [Homo sapiens]	1	184: NLLGG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKYSFSG	FLIFP
1	12: LSNVE	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRKSSYSG	SNGGD	CAB66277	ck: 4543 len: 65	! GB:AL939113 hypothetical protein SCC57A.09 [Homo sapiens]	1	AAF32315	ck: 2510 len: 197	! GB:AF218380 CBLN3 [Mus musculus] (ver 1)
1	7: SIMDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRKSSYSG	PDDGN	CAB66278	ck: 8793 len: 65	! GB:AL939113 hypothetical protein SCC57A.10 [Homo sapiens]	1	30: TALPO	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKIYSYAG	DNIVT
1	7: REMDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRKSSYSG	PDDGN	CAB59668	ck: 2102 len: 519	! GB:AL939115 putative monooxygenase [Streptococcus pneumoniae]	1	AAL87620	ck: 4224 len: 1,084	! GB:AF411442 DLC-1 [Mus musculus] (ver 1)
1	50: SVGGT	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRDNSYSG	CACDW	CAD30920	ck: 2209 len: 266	! GB:AL939121 putative enoyl-CoA hydratase [Streptococcus pneumoniae]	1	394: NSGVN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG WRTGSFHG	PGHLS
1	239: VEDAA	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WESVAFSG	DRAEG	CAC33901	ck: 1282 len: 243	! GB:AL939123 hypothetical protein SCBAC5H2.0 [Homo sapiens]	1	CAC27318	ck: 9148 len: 374	! GB:AX069389 Human class III alcohol dehydrogenase
1	47: HRGHR	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKAASFILG	VDEVY	CAA06915	ck: 9067 len: 432	! GB:AJ006215 CMP-N-acetylneuraminic acid synthetase [Homo sapiens]	1	314: VTGRT	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGTAFGG	WKSVE
1	363: EMGLC	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKEVAIILG	NEVSD	AAF98323	ck: 8989 len: 193	! GB:AF249889 claudin-19 [Mus musculus] (ver 1)	1	30: CALPM	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKVTAFIG	NSIVV
1	20: TALPO	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WQSSSYAG	DAIIT	CAA09727	ck: 8372 len: 219	! GB:AJ011656 Claudin-3 [Rattus norvegicus] (ver 1)	1	30: MLLPS	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKTSSYVG	ASIVT
1	29: CALPM	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WVSAFIFG	SSIIT				1	CAC39830	ck: 587 len: 496	! GB:AX136377 unnamed protein product [Homo sapiens]
							1	4: MAL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRGSAYAG	FLALA
							1	CAC38585	ck: 587 len: 496	! GB:AX119023 unnamed protein product [Homo sapiens]

```

1      W(K,R)XX(S,A)(Y,F)XG
4: MAL  WEGSAVAG  FLALA
CAD20708  ck: 2066  len: 211  ! GB:AX342328 unnamed protein product [Homo s
1      W(K,R)XX(S,A)(Y,F)XG
30: TALPQ  WRIYSYAG  DNIVT
CAC88622  ck: 1615  len: 237  ! GB:AX247812 unnamed protein product [Homo s
1      W(K,R)XX(S,A)(Y,F)XG
4: MAL  WEGSAVAG  FLALA
CAA55971  ck: 2936  len: 159  ! GB:X79465 ORF II [Cauliflower mosaic virus]
1      W(K,R)XX(S,A)(Y,F)XG
61: SLIGI  WKINSYFG  LSKDP
CAB75437  ck: 3583  len: 479  ! GB:AJ249447 putative aphid transmission fac
1      W(K,R)XX(S,A)(Y,F)XG
129: ANLAG  WRAYAYSG  CTISN
AAD37338  ck: 2997  len: 159  ! GB:AF140604 aphid transmission helper facto
1      W(K,R)XX(S,A)(Y,F)XG
61: SLIGI  WKINSYFG  LSKDP
AAK14522  ck: 1050  len: 284  ! GB:AF204951 EsV-1-104 [Ectocarpus siliculos
1      W(K,R)XX(S,A)(Y,F)XG
75: NCLKN  WKFRAPLG  SGAGG
CAB54735  ck: 405  len: 146  ! GB:Y16780 A32L protein [Variola minor virus
1      W(K,R)XX(S,A)(Y,F)XG
73: DVKQK  WRCVAYPG  NGFVS
AAD45687  ck: 7259  len: 479  ! GB:AF157029 putative aphid transmission fac
1      W(K,R)XX(S,A)(Y,F)XG
129: ANLAG  WRAYAYSG  CTISN
AAK37057  ck: 2878  len: 138  ! GB:AF281817 t22.9 [Tupala herpesvirus] (ver
1      W(K,R)XX(S,A)(Y,F)XG
67: TSVPV  WRICSFSG  GFVLG
AAL98880  ck: 8999  len: 861  ! GB:AF383258 envelope glycoprotein [Human im
1      W(K,R)XX(S,A)(Y,F)XG

```

```

777: LASGI  WRVNSYLK  LGLGI
AAG37639  ck: 663  len: 146  ! GB:AY009089 CMP145L [Camelpox virus CMS]
1      W(K,R)XX(S,A)(Y,F)XG
73: DVKQK  WRCVAYPG  NGFVS
AAL73854  ck: 663  len: 146  ! GB:AF438165 hypothetical protein; CMLV14
1      W(K,R)XX(S,A)(Y,F)XG
73: DVKQK  WRCVAYPG  NGFVS
AAP21790  ck: 8531  len: 676  ! GB:AY236971 putative aphid transmission
1      W(K,R)XX(S,A)(Y,F)XG
326: ANLAG  WRAYAYSG  CTISN
AAK13091  ck: 3386  len: 1,641  ! GB:AC078839 Mutator-like transposase [Or
1      W(K,R)XX(S,A)(Y,F)XG
266: KEDCP  WRVHAYKG  KWNDY
AAG48840  ck: 2655  len: 1,011  ! GB:AC084218 similar to Oryza sativa Muta
1      W(K,R)XX(S,A)(Y,F)XG
277: KDGCP  WRVHAYKG  KWKDY
AAK55480  ck: 4734  len: 883  ! GB:AC084295 putative transposase related
1      W(K,R)XX(S,A)(Y,F)XG
266: KEDCP  WRVHAYKG  VEKYH
AAB71967  ck: 2687  len: 428  ! GB:AC002292 putative Clathrin Coat Assem
1      W(K,R)XX(S,A)(Y,F)XG
346: KDALV  WKIKSFPK  NKEYM
AAK62656  ck: 5269  len: 379  ! GB:AY039601 AT5g43940/MRH10_4 [Arabidops
1      W(K,R)XX(S,A)(Y,F)XG
317: VTGRV  WKGTAFGG  FKSRT
AAK72879  ck: 695  len: 2,910  ! GB:AF378327 polyunsaturated fatty acid s
1      W(K,R)XX(S,A)(Y,F)XG
181: ANPPE  WRLDSFPK  FLGNV
AAK98730  ck: 3231  len: 2,421  ! GB:AC090485 Putative mutator-like transp
1      W(K,R)XX(S,A)(Y,F)XG
172: KEDCP  WRVHAYKG  KWNDY

```


1	AAK87310	ck: 5850 len: 298	! GB:AE008075 AGR_c_2802p [Agrobacterium tumefaciens]	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRMAYFG	DTMAH	309: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGSAFVG	VKGRT
1	AAL42524	ck: 9687 len: 272	! GB:AE009110 ABC transporter, membrane spanning	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRRMAYFG	DTMAH	308: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGSAFVG	VKGRS
1	AAK38391	ck: 2507 len: 285	! GB:AF261825 hypothetical protein [Salmonella enterica]	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKTAFVFG	SGNID	374: PSQDG	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKKSAFVG	APAFG
1	AAL94075	ck: 4108 len: 454	! GB:AE010501 Inner membrane protein [Fusobacterium nucleatum]	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG WKSPSFG	VLPTG	AAK43139	ck: 8382 len: 300	! GB:AE012512 integral membrane protein [Xanthomonas campestris]
1	AAL95234	ck: 5980 len: 303	! GB:AE010611 Transporter [Fusobacterium nucleatum]	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKSIAVAG	FLSSG	379: PTPAQ	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKNVAFMG	AALLL
1	AAK07944	ck: 6769 len: 355	! GB:AE011182 iron(III) ABC transporter, permease	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRSIAFAG	FLLLP	374: PSPDG	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKKSAFVG	APSGG
1	AAK25260	ck: 4513 len: 474	! GB:AE013156 S-layer homology domain [Thermotoga maritima]	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG WKLAYDGS	VYNPV	AAK35517	ck: 6719 len: 697	! GB:AE011691 prolyl oligopeptidase [Xanthomonas campestris]
1	AAK25829	ck: 1574 len: 835	! GB:AE013210 conserved hypothetical protein	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKDIAFAG	TLNGA	AAK35623	ck: 6865 len: 369	! GB:AE011703 alcohol dehydrogenase C [Xanthomonas campestris]
1	AAK40941	ck: 898 len: 205	! GB:AE012266 conserved hypothetical protein	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRAFQG	YADHM	309: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGSAFVG	VKGRT
1	AAK42350	ck: 8539 len: 811	! GB:AE012422 ferric enterobactin receptor [Xanthomonas campestris]	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRAFQG	YADHM	AAK36532	ck: 8710 len: 201	! GB:AE011799 conserved hypothetical protein
1	AAK42659	ck: 6765 len: 369	! GB:AE012457 alcohol dehydrogenase C [Xanthomonas campestris]	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRAFQG	DLILL	102: SPNGA	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRNAFQG	YADHM
1	AAK42659	ck: 6765 len: 369	! GB:AE012457 alcohol dehydrogenase C [Xanthomonas campestris]	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRAFQG	DLILL	AAK38052	ck: 333 len: 811	! GB:AE011966 ferric enterobactin receptor
1	AAK42659	ck: 6765 len: 369	! GB:AE012457 alcohol dehydrogenase C [Xanthomonas campestris]	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRAFQG	DLILL	645: ELVAN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRQYAFVG	DLILL
1	AAK42659	ck: 6765 len: 369	! GB:AE012457 alcohol dehydrogenase C [Xanthomonas campestris]	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRAFQG	DLILL	AAK38996	ck: 1130 len: 423	! GB:AE012066 cation efflux system protein
1	AAK42659	ck: 6765 len: 369	! GB:AE012457 alcohol dehydrogenase C [Xanthomonas campestris]	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRAFQG	DLILL	146: GKHR	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRPSPYBG	RTTIA
1	AAK42659	ck: 6765 len: 369	! GB:AE012457 alcohol dehydrogenase C [Xanthomonas campestris]	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRAFQG	DLILL	AAK23122	ck: 9243 len: 613	! GB:AE005792 TonB-dependent receptor, putative
1	AAK42659	ck: 6765 len: 369	! GB:AE012457 alcohol dehydrogenase C [Xanthomonas campestris]	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRAFQG	DLILL	370: KGDTH	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRRAFQG	FRPPT

1	AAK24487	ck: 8539 len: 369	! GB:AE005919 alcohol dehydrogenase, class II	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGTAFFG	ARGRT
1	309: VTGRV				
1	AAH85811	ck: 7586 len: 261	! GB:AE013828 inner membrane permease of zind	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRMSYFG	DTLAH
1	28: GSFVV				
1	AAH86220	ck: 2620 len: 379	! GB:AE013869 alcohol dehydrogenase class III	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WKGSAFAG	VKGRS
1	312: VTGRV				
1	AAH86250	ck: 7447 len: 766	! GB:AE013872 hypothetical [Yersinia pestis K	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTISYFG	GKGVV
1	306: ADAPL				
1	AAH86714	ck: 1453 len: 1,241	! GB:AE013917 DNA helicase, ATP-dependent ds	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRVTSISG	LOQSG
1	925: KMODY				
1	AAH86903	ck: 8384 len: 765	! GB:AE013936 hypothetical [Yersinia pestis K	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTISYFG	GKGVV
1	306: ADAPL				
1	AAK20945	ck: 2066 len: 211	! GB:AF260406 claudin-1 [Homo sapiens] (ver 1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTISYAG	DNIVT
1	30: TALPQ				
1	AAH06112	ck: 3596 len: 497	! GB:BC006112 Similar to RIKEN cDNA 2610017G	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG WRGSAYAG	FLALA
1	4: MAL				
1	AAH08934	ck: 5150 len: 220	! GB:BC008934 claudin 6 [Homo sapiens] (ver 1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKVTAFIG	NSIVV
1	30: CALPM				
1	AAH12471	ck: 2066 len: 211	! GB:BC012471 claudin 1 [Homo sapiens] (ver 1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRTISYAG	DNIVT
1	30: TALPQ				
1	AAK97501	ck: 9918 len: 1,091	! GB:AF408781 deleted in liver cancer 1 [Homo		

1	401: NGSVN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG WKGSAFHG	PGHIS
1	AAH14424	ck: 8102 len: 230	! GB:BC014424 Similar to claudin 2 [Homo
1	30: MLLPS	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKTSSYVG	ASIVT
1	AAH14665	ck: 9148 len: 374	! GB:BC014665 Unknown (protein for MGC:23
1	314: VTGRT	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKGTAFFG	WKSVE
1	AAH16056	ck: 1473 len: 220	! GB:BC016056 claudin 3 [Homo sapiens] (v
1	29: CALPM	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRVSAFAG	SNIIIT
1	CAD90967	ck: 6585 len: 573	! GB:AX741417 unnamed protein product [Es
1	415: SIETD	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGRAFIG	YLSPY
1	CAD90968	ck: 9898 len: 437	! GB:AX741417 unnamed protein product [Es
1	279: SIETD	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGRAFIG	YLSPY
1	CAD90969	ck: 7330 len: 303	! GB:AX741417 unnamed protein product [Es
1	145: SIETD	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGRAFIG	YLSPY
1	CAD90970	ck: 5402 len: 182	! GB:AX741417 unnamed protein product [Es
1	24: SIETD	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRGRAFIG	YLSPY
1	AAK77295	ck: 4884 len: 1,721	! GB:AY047563 GH07636p [Drosophila melanc
1	253: SCMLN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG WRPPSYDG	GLKVS
1	AAK28197	ck: 8082 len: 255	! GB:AF338091 cytochrome b [Blackburnia f
1	85: GYVLP	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG WRQMSFWG	ATVIT
1	AAK26851	ck: 835 len: 377	! GB:AF344170 alcohol dehydrogenase class
1		W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	

317: VTGRV WKGTAFGG WKSVE

AAK26852 ck: 1200 len: 377 ! GB:AF344171 alcohol dehydrogenase class 3 [

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE

317: VTGRV WKGTAFGG

AAL72130 ck: 6247 len: 377 ! GB:AF344172 alcohol dehydrogenase class 3 [

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG YKSVE

317: VTGRV WKGTAFGG

AAL72131 ck: 6358 len: 377 ! GB:AF344173 alcohol dehydrogenase class 3 [

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG YKSVE

317: VTGRV WKGTAFGG

AAL90256 ck: 1032 len: 379 ! GB:AY089518 GM08044p [Drosophila melanogast

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WRSVS

319: VTGRV WKGSAFGG

AAL90353 ck: 1032 len: 379 ! GB:AY089615 RE29421p [Drosophila melanogast

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WRSVS

319: VTGRV WKGSAFGG

AAM26273 ck: 7283 len: 99 ! GB:AF266719 alcohol dehydrogenase class 3 [

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVE

39: VTGRV WKGTAFGG

AAF73255 ck: 745 len: 377 ! GB:AF156705 alcohol dehydrogenase class 3 [

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG WKSVD

317: VTGRV WKGTAFGG

AAK39164 ck: 8316 len: 545 ! GB:U23529 Hypothetical protein C15B12.4 [Ca

1 W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG FSNV

404: GLVGG WENSSFSG

AAA80412 ck: 2841 len: 551 ! GB:U39666 Hypothetical protein K04E7.3 [Cae

1 W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG SSDCY

376: RADYS WENISISG

AAF60807 ck: 7111 len: 317 ! GB:AC024827 Hypothetical protein Y55F3AR.1

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG VLAFF

241: VVKMT WKQKSIDG

1	AAH24178	ck: 6092	len: 643	! GB:BC024178 hypothetical protein FLJ1009
		W(K,R)XX(S,A)(Y,F)XG		
	632: TLEKR	WKLCAFEG	IKIT	
1	AAH26325	ck: 5678	len: 376	! GB:AF399909 alcohol dehydrogenase [Danio
		W(K,R)XX(S,A)(Y,F)XG		
	316: VTGRT	WKGTAFG	WKSVE	
1	AAH37302	ck: 8315	len: 209	! GB:AF323909 E-NTPDase [Torpedo marmorata
		W(K,R)XX(S,A)(Y,F)XG		
	90: QSSCA	WKQCSFNG	YQPH	
1	AAH24512	ck: 4672	len: 215	! GB:AF260240 claudin7 [Danio rerio] (ver
		W(K,R)XX(S,A)(Y,F)XG		
	30: TIMPQ	WKMSAYVG	DNIIT	
1	AAH42293	ck: 4992	len: 211	! GB:BC042293 Similar to claudin 1 [Xenopus
		W(K,R)XX(S,A)(Y,F)XG		
	30: IAIPO	WKMSFAG	DAIIT	
1	AAH48771	ck: 6651	len: 213	! GB:BC048771 Unknown (protein for MGC:534
		W(K,R)XX(S,A)(Y,F)XG		
	30: TIMPQ	WKMSYAG	DQIIT	
1	AAH48367	ck: 4672	len: 215	! GB:BC048367 Unknown (protein for MGC:552
		W(K,R)XX(S,A)(Y,F)XG		
	30: TIMPQ	WKMSAYVG	DNIIT	
1	AAH46020	ck: 7980	len: 199	! GB:BC046020 Cldn1 protein [Danio rerio]
		W(K,R)XX(S,A)(Y,F)XG		
	30: IAIPO	WKTSAFIG	QNIIT	
1	AAH44457	ck: 1250	len: 209	! GB:BC044457 Claudin g [Danio rerio] (ver
		W(K,R)XX(S,A)(Y,F)XG		
	29: CAIPL	WFTVAFIG	NNIVT	
1	AAH53223	ck: 4987	len: 214	! GB:BC053223 Unknown (protein for MGC:640
		W(K,R)XX(S,A)(Y,F)XG		
	29: CALPM	WVNSAFVG	ANIVT	
1	AAH20503	ck: 2687	len: 428	! GB:AY099652 clathrin adaptor medium chain
		W(K,R)XX(S,A)(Y,F)XG		

346: KDALV	W(K,R)XX(S,A)(Y,F)XG WKIKSFPG	NKEYM	AAO37500	ck: 4142	len: 664	! GB:AC092262 putative transposon protein
AAAM15044	ck: 7546	len: 1,557	! GB:AC099774 putative transposase related pr	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRVHAYKG	KWKDY
266: KEDCP	WRVHAYKG	KWKDY	AAO39856	ck: 8691	len: 342	! GB:AC133778 hypothetical protein [Oryza
AAAM94925	ck: 2184	len: 655	! GB:AC091122 mutator-like transposase, 3'-pa	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRVHAYKG	KWKDY
266: KEDCP	WRVHAYKG	KWKDY	AAO73287	ck: 959	len: 1,153	! GB:AC099324 putative mutator-like transp
AAAM95679	ck: 3122	len: 1,536	! GB:AC090870 putative Mu transposable element	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRVHAYKG	KWKDY
266: KEDCP	WRVHAYKG	KWKDY	AAO03385	ck: 3009	len: 1,656	! GB:AC116369 putative mutator-like transp
AAAM05493	ck: 1340	len: 779	! GB:AC099399 Putative transposase protein [O	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRVHAYKG	KWKDY
266: KEDCP	WRVHAYKG	KWKDY	AAO03412	ck: 6702	len: 844	! GB:AC118133 putative mutator-like transp
AAAM61479	ck: 7087	len: 1,596	! GB:AC105733 Putative mutator-like transpos	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRVHAYKG	KWKDY
277: KDGCP	WRVHAYKG	KWKDY	AAAM64806	ck: 5269	len: 379	! GB:AY087250 alcohol dehydrogenase (EC 1
AAAM72155	ck: 2687	len: 428	! GB:BT002144 clathrin adaptor medium chain F	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRVHAYKG	KWKDY
346: KDALV	WKIKSFPG	NKEYM	AAAP12877	ck: 3045	len: 265	! GB:BT006228 At4g38890 [Arabidopsis thal:
AAO19375	ck: 9506	len: 1,527	! GB:AC123974 putative mutator-like transpos	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRVHAYKG	KWKDY
267: KDGCP	WRVHAYKG	KWKDY	AAH30524	ck: 2934	len: 211	! GB:BC030524 similar to claudin 19 [Homo
AAO27752	ck: 8680	len: 230	! GB:AF359360 deacetylase [Fusarium sporotric	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG	WRVHAYKG	KWKDY
170: IDTDD	WRVHAYKG	AEPSL	AAH98151	ck: 8102	len: 230	! GB:AF250558 claudin-2 [Homo sapiens] (v
AAO34677	ck: 8243	len: 230	! GB:AF359361 deacetylase [Gibberella zeae] (W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(F)XG	WRVHAYKG	KWKDY
170: IDTDD	WRVHAYKG	AEPSL	AAH50138	ck: 3262	len: 219	! GB:BC050138 claudin 6 [Mus musculus] (v

1 W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
30: CALPM WKVTAFIG NSIVV

AAH04606 ck: 2584 len: 166 ! GB:BC004606 Cmas protein [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
97: EMGLC WKEVAYLG NEVSD

AAH05718 ck: 3382 len: 219 ! GB:BC005718 claudin 6 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
30: CALPM WKVTAFIG NSIVV

AAH15252 ck: 6271 len: 230 ! GB:BC015252 claudin 2 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
30: MLPN WRTSSYVG ASIVT

AAH12650 ck: 147 len: 200 ! GB:BC012650 Cldn3 protein [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
29: CALPM WRVSAFIG SSIT

AAH23094 ck: 3258 len: 490 ! GB:BC023094 RIKEN CDNA 4932432K03 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
198: OKOPA WKVGAYKG KAQIS

AAH31500 ck: 8878 len: 432 ! GB:BC031500 cytidine monophospho-N-acetylglutamate synthase [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
363: EMGLC WKEVAYLG NEVSD

AAH02003 ck: 2898 len: 211 ! GB:BC002003 Claudin 1 [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
30: TALPQ WKIYSYAG DNIVT

AAH29630 ck: 5911 len: 213 ! GB:AE014376 outer-membrane protein, 25 kDa
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
74: IKPDD WKAGAFAG WNEQQ

AAH34284 ck: 2536 len: 286 ! GB:AE014604 zinc ABC transporter, permease
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
42: GCFII WRRWAYFG DTMAH

AAH43426 ck: 121 len: 253 ! GB:AE015206 orf, conserved hypothetical protein [Mus musculus] (ver 1)
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
*20: GSFVV WRRMSYFG DTLAH

AAH44318 ck: 5059 len: 1,183 ! GB:AE015297 DNA helicase, ATP-dependent

904: LPGEN W(R)XX(S)(Y)XG
WRVTSYSG LQORG

AAH57900 ck: 8469 len: 372 ! GB:AE014863 putative alcohol dehydrogenase

312: VTGRV W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
WRGSAFAG VKGKT

AAH23891 ck: 8641 len: 366 ! GB:AE014616 permease protein of ABC tran

131: LMIFF W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
WRWFAFAG LPSPS

AAH55101 ck: 2100 len: 379 ! GB:AE015647 alcohol dehydrogenase class

317: VTGRV W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
WRGSAFAG VKGRS

AAH55193 ck: 5478 len: 1,259 ! GB:AE015657 exodeoxyribonuclease V, beta

949: QYRFP W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
WRVGSYSG LVKNA

AAH65751 ck: 9475 len: 260 ! GB:AE016774 zinc ABC transporter, permea

28: GSFVV W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
WRRWAYFG DTLSH

AAH67237 ck: 7768 len: 371 ! GB:AE016780 D-isomer specific 2-hydroxya

311: VTGRV W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
WRGSAFAG VRGRS

AAH67553 ck: 8502 len: 978 ! GB:AE016781 hypothetical protein [Pseu

30: LLAEV W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
WRTASYLG SNYAP

AAH68164 ck: 2550 len: 226 ! GB:AE016783 conserved hypothetical prote

166: DYQDG W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
WRRESFAG FAGDW

AAH78842 ck: 7207 len: 715 ! GB:AE016756 Putative cytoplasmic membran

148: FWRVL W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(F)XG
WRFRAYTG QVIVA

AAH78943 ck: 8529 len: 369 ! GB:AE016756 Alcohol dehydrogenase class

1	309: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKGSAFGG	VKGRS	
1	AAN80056	ck: 9281 len: 247	! GB:AE016759 Putative tail component of pro		
1	228: WEHRA	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG	WRASAFG	IYNDL	
1	AAN80730	ck: 7660 len: 261	! GB:AE016761 High-affinity zinc uptake syste		
1	28: GSFVV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	WRMSYFG	DTLAH	
1	AAN80978	ck: 9474 len: 261	! GB:AE016762 Hypothetical protein [Escherich		
1	71: DEIKY	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	WRYTSYKG	GEFPE	
1	AAN81859	ck: 4588 len: 1,183	! GB:AE016765 Exodeoxyribonuclease V beta cha		
1	904: LPGDN	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	WRVTSYSG	LQQRG	
1	AAN53077	ck: 4584 len: 376	! GB:AE015922 zinc-binding dehydrogenase [She		
1	314: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKGSAFGG	VKGRS	
1	AA035232	ck: 1188 len: 241	! GB:AE015938 conserved protein [Clostridium		
1	191: NNNKL	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKIDAFG	VMFLA	
1	AAM29698	ck: 4651 len: 535	! GB:AE013219 Dipeptide ABC transporter, bind		
1	469: QAIEN	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	WKLADYDG	NTCFG	
1	AA050721	ck: 8784 len: 538	! GB:AE004659 probable sulfatase [Pseudomonas		
1	455: VRSER	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(Y)XG	WKYIADG	FRAQL	
1	AA07017	ck: 6470 len: 370	! GB:AE004783 alcohol dehydrogenase class III		
1	310: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG	WRGSAFGG	VRGRS	
1	AAG08986	ck: 5958 len: 262	! GB:AE004962 permease of ABC zinc transporte		
1	310: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG			

28: GSFVV	WRRMAYFG	DTLSH			
AA008691	ck: 1793 len: 265	! GB:AE016797 ABC-type Mn2+/Zn2+ transpor			
32: GSFVV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRRMAYFG	DTLAH		
AA008871	ck: 391 len: 376	! GB:AE016798 Zn-dependent alcohol dehydr			
315: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG	WRGSAFGG	VKGRS		
AA010204	ck: 5828 len: 1,206	! GB:AE016802 ATP-dependent exoDNase beta			
909: PIDRL	W(K,R)XX(S,A)(Y,F)XG W(R)XX(S)(Y)XG	WRITSYSG	LVKQG		
AA054357	ck: 7617 len: 1,067	! GB:AE016858 type IV pilus-associated pr			
679: AAGSS	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRTVAYGG	TGAGG		
AA055078	ck: 7207 len: 370	! GB:AE016861 alcohol dehydrogenase, clas			
310: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG	WRGSAFGG	VRGRT		
AA055981	ck: 2136 len: 316	! GB:AE016864 DNA/RNA non-specific endonu			
231: VPLAF	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKVAVFVG	DDGKP		
AA058691	ck: 7040 len: 262	! GB:AE016875 zinc ABC transporter, perme			
28: GSFVV	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG	WRRMAYFG	DTLSH		
AAC07799	ck: 5280 len: 392	! GB:AE000769 hypothetical protein [Aquil			
127: YIVSVN	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(F)XG	WKERSFEG	KDFED		
AA152430	ck: 5809 len: 213	! GB:AE009564 25 KDA OUTER-MEMBRANE IMMUN			
74: IXPDD	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WRAGAFAG	WNFQQ		
AA153000	ck: 9722 len: 370	! GB:AE009614 glutathione-dependent form			
310: VTGRV	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG	WKGTAFGG	ARGRT		

1 AAL53417 ck: 8470 len: 284 ! GB:AE009656 HIGH-AFFINITY ZINC UPTAKE SYSTEM
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(A)(Y)XG
40: GCFII WRRMAIFG DTMAR

1 AAO68656 ck: 7341 len: 261 ! GB:AE016837 high-affinity zinc uptake system
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
28: GSFVV WRRMSYFG DTLAH

1 AAO68933 ck: 4455 len: 122 ! GB:AE016838 putative pathogenicity island P
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
70: KALAN WKPAAFOG IPQRL

1 AAO70456 ck: 4255 len: 1,181 ! GB:AE016843 exonuclease V subunit [Salmonella]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
901: LLYDS WRVTSYSG LQQRG

1 AAO76884 ck: 2570 len: 1,019 ! GB:AE016933 conserved hypothetical protein
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
938: ALPDA WKNGSFPG MKARG

1 AAO78593 ck: 7620 len: 561 ! GB:AE016940 arylsulfatase (aryl-sulfate sul
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
404: TPFRE WKAQSYEG GICTP

1 AAO82547 ck: 7654 len: 393 ! GB:AE016955 site-specific recombinase, phag
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
15: KGKXY WKVTAIFG VDYL

1 AAP18143 ck: 1826 len: 1,180 ! GB:AE016987 DNA helicase, ATP-dependent ds
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
901: LPQDN WRVTSYSG LQQRG

1 AAL21870 ck: 4862 len: 1,181 ! GB:AE008837 exonuclease V, beta chain [Salmo
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
901: LLYDS WRVTSYSG LQQRG

1 AAL20337 ck: 5313 len: 122 ! GB:AE008761 secretion system apparatus prot
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(F)XG
70: KALAN WKPAAFOG IPQRL

1 AAL20545 ck: 2185 len: 372 ! GB:AE008771 alcohol dehydrogenase class III
W(K,R)XX(S,A)(Y,F)XG

309: VTGRV W(R)XX(A)(F)XG
WRRSAFEG VKGR

AAL20809 ck: 6897 len: 261 ! GB:AE008784 ABC superfamily high affinity
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
28: GSFVV WRRMSYFG DTLAH

AAP17250 ck: 7607 len: 261 ! GB:AE016984 hypothetical protein [Shigella]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
28: GSFVV WRRMSYFG DTLAH

AA051851 ck: 4765 len: 249 ! GB:AE011616 probable sugar aldolase [Leptospira]
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(Y)XG
211: KYMFS WKSFSYSG EPRKL

AA048391 ck: 7773 len: 523 ! GB:AE011301 putative outer membrane prote
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
136: VTDFI WRLGFSFG EMANR

AA049213 ck: 7773 len: 458 ! GB:AE011374 inner membrane protein CreD-
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
244: QISSN WKDPSFEG SFLPK

AA049861 ck: 8001 len: 749 ! GB:AE011433 ribonuclease BN ribn [Leptospira]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
284: LESPI WRLASFEG VELKD

AA051364 ck: 6380 len: 403 ! GB:AE011570 argininosuccinate synthase [Leptospira]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(F)XG
149: TIAP WRIWSFEG RSDLI

AAP29336 ck: 3946 len: 343 ! GB:AE017041 membrane protein, putative [Leptospira]
W(K,R)XX(S,A)(Y,F)XG
W(R)XX(S)(Y)XG
207: DISKK WRRISYIG GFVGF

AAP11686 ck: 9564 len: 225 ! GB:AE017013 hydrolase (HAD superfamily)
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(S)(F)XG
191: GILGV WKKDSFEG DFKHS

AA049827 ck: 3565 len: 409 ! GB:AY144490 putative fimbrial-like prote
W(K,R)XX(S,A)(Y,F)XG
W(K)XX(A)(Y)XG
180: DRATN WKTYAYGG IVLWI

1	AA057024	ck: 5476	len: 1,274	! GB:AE016970 conserved hypothetical [Mycopla	W(K,R)XX(S,A)(Y,F)XG W(K)XX(S)(Y)XG WKALSYRG	LKERE
80:	FVNLI	WKINSYFG	FKNKE			
1	AA021486	ck: 8280	len: 199	! GB:AE008818 Gifsy-1 prophage protein [Salm	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG WRASAFYG	GLKVS
180:	WRHRA	WRASAFYG	ICNDF			
1	AA044586	ck: 8285	len: 1,214	! GB:BC044586 DEAD/H (Asp-Glu-Ala-Asp/His) bc	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKKHAFGG	WRSVS
532:	NDIPE	WKKHAFGG	NKASY			
1	AA047327	ck: 8168	len: 1,214	! GB:BC047327 DEAD/H (Asp-Glu-Ala-Asp/His) bc	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG WKKHAFGG	GIDNS
532:	NDIPE	WKKHAFGG	NKASY			
1	AA000671	ck: 6154	len: 209	! GB:BT000671 Claudin 4 [Homo sapiens] (ver 1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG W(R)XX(A)(F)XG	GIDNS
30:	CALPM	W(R)XX(A)(F)XG	SNIVT			
1	AA035635	ck: 6154	len: 209	! GB:BT006989 claudin 4 [Homo sapiens] (ver 1	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(F)XG W(R)XX(A)(F)XG	WKSVE
30:	CALPM	W(R)XX(A)(F)XG	SNIVT			
1	AA036063	ck: 5150	len: 220	! GB:BT007399 claudin 6 [Homo sapiens] (ver 1	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG W(K)XX(A)(F)XG	KWKDY
30:	CALPM	W(K)XX(A)(F)XG	NSIVV			
1	AA039899	ck: 9402	len: 384	! GB:AC006663 Hypothetical protein H24K24.3a	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG W(K)XX(A)(F)XG	KWNDY
321:	VTGRT	W(K)XX(A)(F)XG	WKSVE			
1	AA032231	ck: 7595	len: 386	! GB:AC006663 Hypothetical protein H24K24.3b	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG W(K)XX(A)(F)XG	KWNDY
323:	VTGRT	W(K)XX(A)(F)XG	WKSVE			
1	AA021424	ck: 3382	len: 173	! GB:AC006663 Hypothetical protein H24K24.3c	W(K,R)XX(S,A)(Y,F)XG W(K)XX(A)(F)XG W(K)XX(A)(F)XG	KWNDY
110:	VTGRT	W(K)XX(A)(F)XG	WKSVE			
1	AA026360	ck: 9049	len: 1,057	! GB:AE014848 NIMA-related protein kinase (P	W(K,R)XX(S,A)(Y,F)XG W(R)XX(A)(Y)XG W(R)XX(A)(Y)XG	KWNDY
266:	KEDCP	W(R)XX(A)(Y)XG	KWNDY			

```

AAP53739  ck: 3427  len: 865   ! GB:AE017093 unknown protein [Oryza sativa]
1      W(K,R)XX(S,A)(Y,F)XG
      W(R)XX(A)(Y)XG
      264: KDGCP      WRVHAYG      KWNDY

AAP53819  ck: 989   len: 1,421 ! GB:AE017095 putative mutator-like transposase
1      W(K,R)XX(S,A)(Y,F)XG
      W(R)XX(A)(Y)XG
      204: KEDCP      WRVHAYG      KWNDY

AAP53990  ck: 839   len: 1,656 ! GB:AE017099 putative mutator protein [Oryza
1      W(K,R)XX(S,A)(Y,F)XG
      W(R)XX(A)(Y)XG
      266: KEDCP      WRVHAYG      KWNDY

AAP54160  ck: 6567  len: 1,597 ! GB:AE017102 putative transposase [Oryza sat
1      W(K,R)XX(S,A)(Y,F)XG
      W(R)XX(A)(Y)XG
      266: KEDCP      WRVHAYG      KWNDY

AAP54539  ck: 3122  len: 1,536 ! GB:AE017110 putative Mu transposable elemen
1      W(K,R)XX(S,A)(Y,F)XG
      W(R)XX(A)(Y)XG
      266: KEDCP      WRVHAYG      KWNDY

AAP54600  ck: 4017  len: 1,626 ! GB:AE017112 mutator-like transposase [Oryza
1      W(K,R)XX(S,A)(Y,F)XG
      W(R)XX(A)(Y)XG
      266: KEDCP      WRVHAYG      KWNDY

AAP55153  ck: 8786  len: 1,638 ! GB:AE017122 mutator-like transposase [Oryza
1      W(K,R)XX(S,A)(Y,F)XG
      W(R)XX(A)(Y)XG
      266: KEDCP      WRVHAYG      KWNDY

AAH52062  ck: 9624  len: 662   ! GB:BC052062 D230005D02Rik protein [Mus musc
1      W(K,R)XX(S,A)(Y,F)XG
      W(R)XX(A)(F)XG
      651: TLEKR      WKLCAFEG      IKTT

```

Databases searched:

GENPEPT, Release 136.0, Released on 17Jun2003, Formatted on 17Jun2003

```

Total finds:      529
Total length:    449,659,824
Total sequences: 1,453,955
CPU time:        10:51.19

```



```
Script started on Sat Aug 16 14:39:52 2003
stic21% m seq1-pen.res.
> O <
O| |O IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

-- Outline of search "seq1-pen" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "new.key":
seq1 (AA) ID seq1 AA preliminary pattern
1
2 w
2 k or r
2 any character
2 any character
2 s or a
2 y or f
2 any character
2 g

Selected data banks and files:
Data bank : Pending_AA , all entries

-- Output Parameters --

Format Options:
Nucleic acid code matching Exact Indirect file No
Find non-matching hits only No Sequence or key file No
Report key used Yes List of hits No
Note position of hit Yes Hit display Yes
Display full annotations Yes Name and annotations Yes
Sequence context 20

File Options:
Indirect file No
Sequence or key file No
List of hits No
Hit display Yes
Name and annotations Yes

Run mode Batch
Time to start comparison now
Notify at end of run No

-- Run Parameters --

1 match found in sequence:
PCT-US00-05325-2 ; Sequence 2, Application PC/TUS0005325A
(from "/srch/paa/PCTUS_COMB.pep")
Sequence 2, Application PC/TUS0005325A
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Human Glycosylation Enzymes
FILE REFERENCE: PF05PCT
CURRENT APPLICATION NUMBER: PCT/US00/05325A
CURRENT FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/122,409
PRIOR FILING DATE: 1999-03-02
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 434
TYPE: PPT
[7m--More--(0%) [m ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: 133
OTHER INFORMATION: Xaa equals any amino acid
NAME/KEY: SITE
LOCATION: 136
OTHER INFORMATION: Xaa equals any amino acid
NAME/KEY: SITE
```

```
LOCATION: 147
OTHER INFORMATION: Xaa equals any amino acid
NAME/KEY: SITE
LOCATION: 169
OTHER INFORMATION: Xaa equals any amino acid
Found using 'seq1' (new.key)
...

345 VSVSDKLAVDDEWKENGCLWKEVAYLGNEVSDDECLKRVGLSGAPAD
365 372
-----|
1 match found in sequence:
PCT-US00-05882-850 ; Sequence 850, Application PC/TUS0005882
(from "/srch/paa/PCTUS_COMB.pep")
Sequence 850, Application PC/TUS0005882
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben
TITLE OF INVENTION: Human Cancer Associated Gene Sequences and Polypeptides
FILE REFERENCE: PAL06PCT
CURRENT APPLICATION NUMBER: PCT/US00/05882
CURRENT FILING DATE: 2000-03-08
EARLIER APPLICATION NUMBER: 60/124,270
EARLIER FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 850
LENGTH: 383
TYPE: PPT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (299)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
Found using 'seq1' (new.key)
...

303 VAASGEIATRFPLVTCRWKGTAFGGWKSVEVPKLVSVMKKIK
323 330
-----|
1 match found in sequence:
PCT-US00-05882-1376 ; Sequence 1376, Application PC/TUS0005882
(from "/srch/paa/PCTUS_COMB.pep")
[7m--More--(0%) [m [Kstic21%
stic21%
stic21% tail seq1-pen.res
-- Search Statistics --
Times: CPU 00:24:10.08 Total Elapsed 00:26:18.00
Number of sequences searched: 5678955
Number of sequence hits: 3322
Number of separate matches: 3322
Number of sequence hits saved: 0
stic21% ^D
script done on Sat Aug 16 14:40:13 2003
```

```

ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 960296.93511
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
Found using 'seq1' (new.key)

-----
296 VAPAGAEQTTPFLVTVGRVWVGFAGRGRTDVPKIVDWMEGKIQ 316 323
-----
1 match found in sequence:
US-08-922-182-2 : Sequence 2, Application US/08922182
(from "srich/1aa/5B.COMB.pep")
Sequence 2, Application US/08922182
Patent No. 5834300
GENERAL INFORMATION:
APPLICANT: Donohue, Timothy J
APPLICANT: Barber, Robert D
APPLICANT: Witthuhn, Verbon
TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
TITLE OF INVENTION: SENSING AND REMEDIATION
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/922,182
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/608,241
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 960296.93511
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
Found using 'seq1' (new.key)

```

296 VAPAGAEIOTRPFQVYGRVWKGSAFGGARGTDVPKIVDWMKGKIO
316 323

...

1 match found in sequence:
US-08-919-933-2 ; Sequence 2, Application US/08919953
(from "/srch/iaa/5B.COMB.pep")
Sequence 2, Application US/08919953
Patent No. 5837481
GENERAL INFORMATION:
APPLICANT: Donohue, Timothy J
APPLICANT: Barber, Robert D
APPLICANT: Witthuhn, Veiron
TITLE OF INVENTION: MICROBIAL SYSTEM FOR FORMALDEHYDE
TITLE OF INVENTION: SENSING AND REMEDIATION
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/919,953
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/608,241
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 960296.93511
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
Found using 'seq1' (new.key)

...

296 VAPAGAEIOTRPFQVYGRVWKGSAFGGARGTDVPKIVDWMKGKIO
316 323

...

1 match found in sequence:
US-08-637-759B-89 ; Sequence 89, Application US/08637759B
(from "/srch/iaa/5B.COMB.pep")
Sequence 89, Application US/08637759B
Patent No. 5876931
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center

STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,759B
FILING DATE: 03-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
Found using 'seq1' (new.key)

...

10 COILKQTFLELLFKALANKPKAAFGQIPQRLFLLRDGLAMSCSPPL
30 37

...

1 match found in sequence:
US-08-871-355A-89 ; Sequence 89, Application US/08871355A
(from "/srch/iaa/6A.COMB.pep")
Sequence 89, Application US/08871355A
Patent No. 6015669
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,355A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Pabst, Patrea L.
 REGISTRATION NUMBER: 31,284
 REFERENCE/DOCKET NUMBER: RPMS 101 CON
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (404) 873-8794
 TELEFAX: (404) 873-8795
 INFORMATION FOR SEQ ID NO: 89:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 759 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 Found using 'seq1' (new.key)

...
 10 CQILKQFLDELLEKLANWKPAQFQIPQRLFLLRDGIAMSCSPPL
 30 37

 1 match found in sequence:
 US-09-222-817-2 ; Sequence 2, Application US/09222817
 (from "/srch/iaa/6A.COMB.pep")
 Sequence 2, Application US/09222817
 Patent No. 6037154
 GENERAL INFORMATION:
 APPLICANT: Mikiko SUGA, Masakazu SUGIMOTO, Tsuyoshi OSUMI, Tsuyoshi NAKAMATSU,
 TITLE OF INVENTION: METHOD OF PRODUCING L-SERINE BY FERMENTATION
 FILE REFERENCE: OP813
 CURRENT APPLICATION NUMBER: US/09/222,817
 CURRENT FILING DATE: 1998-12-30
 EARLIER APPLICATION NUMBER: JP 10-3751
 EARLIER FILING DATE: 1998-01-12
 EARLIER APPLICATION NUMBER: JP 10-353521
 EARLIER FILING DATE: 1998-12-11
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 2
 LENGTH: 345
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
 Found using 'seq1' (new.key)

...
 112 LLLSTARQIPAADATLREGWKRSFNGVEIFGKTGIVGFGHIGOLF
 132 139

 1 match found in sequence:
 US-09-222-817-12 ; Sequence 12, Application US/09222817
 (from "/srch/iaa/6A.COMB.pep")
 Sequence 12, Application US/09222817
 Patent No. 6037154
 GENERAL INFORMATION:
 APPLICANT: Mikiko SUGA, Masakazu SUGIMOTO, Tsuyoshi OSUMI, Tsuyoshi NAKAMATSU,
 TITLE OF INVENTION: METHOD OF PRODUCING L-SERINE BY FERMENTATION
 FILE REFERENCE: OP813
 CURRENT APPLICATION NUMBER: US/09/222,817
 CURRENT FILING DATE: 1998-12-30
 EARLIER APPLICATION NUMBER: JP 10-3751
 EARLIER FILING DATE: 1998-01-12
 EARLIER APPLICATION NUMBER: JP 10-353521

EARLIER FILING DATE: 1998-12-11
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 12
 LENGTH: 530
 TYPE: PRT
 ORGANISM: Brevibacterium flavum
 Found using 'seq1' (new.key)

...
 112 LLLSTARQIPAADATLREGWKRSFNGVEIFGKTGIVGFGHIGOLF
 132 139

 1 match found in sequence:
 US-09-222-817-14 ; Sequence 14, Application US/09222817
 (from "/srch/iaa/6A.COMB.pep")
 Sequence 14, Application US/09222817
 Patent No. 6037154
 GENERAL INFORMATION:
 APPLICANT: Mikiko SUGA, Masakazu SUGIMOTO, Tsuyoshi OSUMI, Tsuyoshi NAKAMATSU,
 TITLE OF INVENTION: METHOD OF PRODUCING L-SERINE BY FERMENTATION
 FILE REFERENCE: OP813
 CURRENT APPLICATION NUMBER: US/09/222,817
 CURRENT FILING DATE: 1998-12-30
 EARLIER APPLICATION NUMBER: JP 10-3751
 EARLIER FILING DATE: 1998-01-12
 EARLIER APPLICATION NUMBER: JP 10-353521
 EARLIER FILING DATE: 1998-12-11
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 14
 LENGTH: 530
 TYPE: PRT
 ORGANISM: Brevibacterium flavum
 Found using 'seq1' (new.key)

...
 112 LLLSTARQIPAADATLREGWKRSFNGVEIFGKTGIVGFGHIGOLF
 132 139

 1 match found in sequence:
 US-09-192-983-2 ; Sequence 2, Application US/09192983A
 (from "/srch/iaa/6A.COMB.pep")
 Sequence 2, Application US/09192983A
 Patent No. 6242244
 GENERAL INFORMATION:
 APPLICANT: Donohue, Timothy
 APPLICANT: Barber, Robert
 APPLICANT: Witthuhn, Vernon
 TITLE OF INVENTION: Microbial System for Formaldehyde Sensing and
 TITLE OF INVENTION: Remediation
 FILE REFERENCE: 960296.95505
 CURRENT APPLICATION NUMBER: US/09/192,983A
 CURRENT FILING DATE: 1998-11-16
 EARLIER APPLICATION NUMBER: 08/919,953
 EARLIER FILING DATE: 1997-08-29
 EARLIER APPLICATION NUMBER: 08/608,241
 EARLIER FILING DATE: 1996-02-28
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 2
 LENGTH: 376
 TYPE: PRT

ORGANISM: Rhodobacter sphaeroides
Found using 'seq1' (new.key)

...

296 VAPAGAEIOTRPQLVTGVRWKGSAFGARGRTDVPKIVDWYNEGKIQ
316 323

1 match found in sequence:
US-09-222-786-2 ; Sequence 2, Application US/09222786A
(from "/srch/iaa/6A.COMB.pep")
Sequence 2, Application US/09222786A
Patent No. 6258573
GENERAL INFORMATION:
APPLICANT: Mikiko SUGA, Masakazu SUGIMOTO, Tsuyoshi OSUMI, Tsuyoshi NAKAMATSU,
TITLE OF INVENTION: METHOD OF PRODUCING L-SERINE BY FERMENTATION
FILE REFERENCE: OP812
CURRENT APPLICATION NUMBER: US/09/222,786A
CURRENT FILING DATE: 1998-12-30
EARLIER APPLICATION NUMBER: JP 10-3751
EARLIER FILING DATE: 1998-01-12
EARLIER APPLICATION NUMBER: JP 10-353513
EARLIER FILING DATE: 1998-12-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 345
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
Found using 'seq1' (new.key)

...

112 LLLSTARQIPAAATLREGWKRSFNGVEIFGKTGIVGFGHGOLF
132 139

1 match found in sequence:
US-09-222-786-12 ; Sequence 12, Application US/09222786A
(from "/srch/iaa/6A.COMB.pep")
Sequence 12, Application US/09222786A
Patent No. 6258573
GENERAL INFORMATION:
APPLICANT: Mikiko SUGA, Masakazu SUGIMOTO, Tsuyoshi OSUMI, Tsuyoshi NAKAMATSU,
TITLE OF INVENTION: METHOD OF PRODUCING L-SERINE BY FERMENTATION
FILE REFERENCE: OP812
CURRENT APPLICATION NUMBER: US/09/222,786A
CURRENT FILING DATE: 1998-12-30
EARLIER APPLICATION NUMBER: JP 10-3751
EARLIER FILING DATE: 1998-01-12
EARLIER APPLICATION NUMBER: JP 10-353513
EARLIER FILING DATE: 1998-12-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 530
TYPE: PRT
ORGANISM: Brevibacterium flavum
Found using 'seq1' (new.key)

...

112 LLLSTARQIPAAATLREGWKRSFNGVEIFGKTGIVGFGHGOLF
132 139

...

1 match found in sequence:
US-09-222-786-14 ; Sequence 14, Application US/09222786A
(from "/srch/iaa/6A.COMB.pep")
Sequence 14, Application US/09222786A
Patent No. 6258573
GENERAL INFORMATION:
APPLICANT: Mikiko SUGA, Masakazu SUGIMOTO, Tsuyoshi OSUMI, Tsuyoshi NAKAMATSU,
TITLE OF INVENTION: METHOD OF PRODUCING L-SERINE BY FERMENTATION
FILE REFERENCE: OP812
CURRENT APPLICATION NUMBER: US/09/222,786A
CURRENT FILING DATE: 1998-12-30
EARLIER APPLICATION NUMBER: JP 10-3751
EARLIER FILING DATE: 1998-01-12
EARLIER APPLICATION NUMBER: JP 10-353513
EARLIER FILING DATE: 1998-12-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 530
TYPE: PRT
ORGANISM: Brevibacterium flavum
Found using 'seq1' (new.key)

...

112 LLLSTARQIPAAATLREGWKRSFNGVEIFGKTGIVGFGHGOLF
132 139

1 match found in sequence:
US-09-516-143A-2 ; Sequence 2, Application US/09516143A
(from "/srch/iaa/6B.COMB.pep")
Sequence 2, Application US/09516143A
Patent No. 6333182
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Human Glycosylation Enzymes
FILE REFERENCE: PF505PCT
CURRENT APPLICATION NUMBER: US/09/516,143A
CURRENT FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/122,409
PRIOR FILING DATE: 1999-03-02
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 434
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: 133
OTHER INFORMATION: Xaa equals Tyr or His
NAME/KEY: SITE
LOCATION: 136
OTHER INFORMATION: Xaa equals Gly or Val
NAME/KEY: SITE
LOCATION: 147
OTHER INFORMATION: Xaa equals Ser or Pro
NAME/KEY: SITE
LOCATION: 169
OTHER INFORMATION: Xaa equals Gly or Val
Found using 'seq1' (new.key)

...

345 VSVSDKLAVVDENKEMGLCKWEKAYLGNVSDDECLKRVGLSGAPAD

```
...
365 372
-----
1 match found in sequence:
US-09-201-945-89 ; Sequence 89, Application US/09201945
(from "/srch/iaa/6B_COMB.pep")
Sequence 89, Application US/09201945
Patent No. 6342215
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/637,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
Found using 'seq1' (new.key)
...
10 CQILKQTFILDELLFKALANKPAFGQIPQRLFLRLDGLAMSCSPPL
30 37
-----
1 match found in sequence:
US-09-347-878-54 ; Sequence 54, Application US/09347878C
(from "/srch/iaa/6B_COMB.pep")
Sequence 54, Application US/09347878C
Patent No. 6376210
GENERAL INFORMATION:
APPLICANT: Yuan, Chong
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
FILE REFERENCE: 25685-1651
CURRENT APPLICATION NUMBER: US/09/347,878C
CURRENT FILING DATE: 1999-07-06
NUMBER OF SEQ ID NOS: 75
-----
1 match found in sequence:
US-09-324-541-11 ; Sequence 11, Application US/09324541
(from "/srch/iaa/6B_COMB.pep")
Sequence 11, Application US/09324541
Patent No. 6391855
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
TITLE OF INVENTION: JUNCTIONAL ADHESION MOLECULE-MEDIATED FUNCTIONS
FILE REFERENCE: 100086.412
CURRENT APPLICATION NUMBER: US/09/324,541
CURRENT FILING DATE: 1999-06-02
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 8
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fab fragment directed against claudin cell
OTHER INFORMATION: adhesion recognition sequence
Found using 'seq1' (new.key)
1 WKIYSYAG
1
-----
1 match found in sequence:
US-09-130-491-4 ; Sequence 4, Application US/09130491
(from "/srch/iaa/6B_COMB.pep")
Sequence 4, Application US/09130491
Patent No. 6416974
GENERAL INFORMATION:
APPLICANT: Holtzman, Douglas A.
APPLICANT: Godegari, Andrew D.J.
TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
FILE REFERENCE: 09404/041001
CURRENT APPLICATION NUMBER: US/09/130,491
CURRENT FILING DATE: 1998-08-07
EARLIER APPLICATION NUMBER: US 60/058,108
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: US 60/054,961
EARLIER FILING DATE: 1997-08-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
Found using 'seq1' (new.key)
...
10 GFILAFLGWIGAIIVSTALPQWRISYAGDNIVTAQAMIEGLWNSCVSQ
|-----|
```

30 37

...

1 match found in sequence:

US-09-130-491-15 ; Sequence 15, Application US/09130491
(from "/srch/iaa/6B_COMB.pep")
Sequence 15, Application US/09130491
Patent No. 6416974

GENERAL INFORMATION:

APPLICANT: Holtzman, Douglas A.
APPLICANT: Goodearl, Andrew D.J.
TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
FILE REFERENCE: 09404/041001
CURRENT APPLICATION NUMBER: US/09/130,491
CURRENT FILING DATE: 1998-08-07
EARLIER APPLICATION NUMBER: US 60/058,108
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: US 60/054,961
EARLIER FILING DATE: 1997-08-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15

LENGTH: 247

TYPE: PRT

ORGANISM: Homo sapiens

Found using 'seq1' (new.key)

...

29 36

9 GTSLVGLVGLCTIVCCALPMVRSAFTGSSIIITQAQTWEGLWMNCVQS
29 36

...

1 match found in sequence:

US-09-724-623-110 ; Sequence 110, Application US/09724623
(from "/srch/iaa/6B_COMB.pep")
Sequence 110, Application US/09724623
Patent No. 6476209

GENERAL INFORMATION:

APPLICANT: Glenn, Matthew
APPLICANT: Lubbers, Mark W
APPLICANT: Dekker, James
TITLE OF INVENTION: Polynucleotides, materials incorporating
FILE OF INVENTION: them, and methods for using them.

FILE REFERENCE: 1048U1

CURRENT APPLICATION NUMBER: US/09/724,623

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 124

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 110

LENGTH: 306

TYPE: PRT

ORGANISM: Lactobacillus rhamnosus

Found using 'seq1' (new.key)

...

179 186

159 NKSLNQEATFVSFGPDTRWRRTATSGLFNDNDKPLRDYTPFEMKLLL
179 186

...

1 match found in sequence:

US-08-976-063E-8 ; Sequence 8, Application US/08976063E
(from "/srch/iaa/6B_COMB.pep")
Sequence 8, Application US/08976063E

EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 344
LENGTH: 202
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (202)
OTHER INFORMATION: xaa equals stop translation
Found using 'seq1' (new.key)

...

2 GIALAVLGWLVMLCCALPMWRVTAFIGSNIVTSQTIWEGLNWNCVQ
22 29

...

1 match found in sequence:

US-09-205-258-463 ; Sequence 463, Application US/09205258
(from "/src/iaa/6B.COMB.pep")
Sequence 463, Application US/09205258

Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: PZ007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06


```
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 463
LENGTH: 206
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (80)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (206)
OTHER INFORMATION: Xaa equals stop translation
Found using 'seq1' (new.key)
...
172  PLDPGRVSLRLRNLGGKYSFSGFLIPFLX
      |-----|
      192 199
1 match found in sequence:
US-09-252-991A-18795 : Sequence 18795, Application US/09252991A
(from "/srch/iaa/65.COMB.pep")
Sequence 18795, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18795
LENGTH: 334
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
Found using 'seq1' (new.key)
...
80  ALLAGLALVAGPLGSEVWVRMYATGDTLSHAALLGLVGLGFLDVS
      |-----|
      100 107
1 match found in sequence:
US-09-252-991A-18795 : Sequence 18795, Application US/09252991A
(from "/srch/iaa/65.COMB.pep")
Sequence 18795, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18795
LENGTH: 334
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
Found using 'seq1' (new.key)
...
1 match found in sequence:
US-09-252-991A-18795 : Sequence 18795, Application US/09252991A
(from "/srch/iaa/65.COMB.pep")
Sequence 18795, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18795
LENGTH: 334
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
Found using 'seq1' (new.key)
...
1 match found in sequence:
US-09-252-991A-22187 : Sequence 22187, Application US/09252991A
(from "/srch/iaa/6B.COMB.pep")
Sequence 22187, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
```

```
1 match found in sequence:
US-09-252-991A-20880 : Sequence 20880, Application US/09252991A
(from "/srch/iaa/6B.COMB.pep")
Sequence 20880, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20880
LENGTH: 375
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
Found using 'seq1' (new.key)
...
335  AAGVGQGTTHORSWFIRSITWRVMSYCGLTSTTAPRSMTWS
      |-----|
      355 362
1 match found in sequence:
US-09-252-991A-21227 : Sequence 21227, Application US/09252991A
(from "/srch/iaa/6B.COMB.pep")
Sequence 21227, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 21227
LENGTH: 399
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
Found using 'seq1' (new.key)
...
319  VAGAQEISTRPQLVTGVRGSRFPGYGRGSELPYSYVEKAKGKEIP
      |-----|
      339 346
1 match found in sequence:
US-09-252-991A-22187 : Sequence 22187, Application US/09252991A
(from "/srch/iaa/6B.COMB.pep")
Sequence 22187, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
```

CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 22187
 LENGTH: 728
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa

FEATURE:
 NAME/KEY: UNSURE
 LOCATION: (699)
 OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
 Found using 'seq1' (new.key)

...

7 ALSRGIRTVRTARGPQAQWIPAPFGAACKGLPGRGRRGAGRLV
 27 34

...

1 match found in sequence:
 US-09-252-991A-25507 ; Sequence 25507, Application US/09252991A
 (from "/srch/iaa/6B.COMB.pep")
 Sequence 25507, Application US/09252991A
 Patent No. 6551795

GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 25507
 LENGTH: 447

TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 Found using 'seq1' (new.key)

...

53 RYSLTPTASRASAGSVGWRRAAFIGVSTPLQRRRAVLSEIVLPGR
 73 80

...

1 match found in sequence:
 US-09-252-991A-25705 ; Sequence 25705, Application US/09252991A
 (from "/srch/iaa/6B.COMB.pep")
 Sequence 25705, Application US/09252991A
 Patent No. 6551795

GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 25705
 LENGTH: 547
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 Found using 'seq1' (new.key)

...

444 RERLGRPIDACRMVMRSEWKYIAYDGFRAQLFDLASDPGEIRDIGA
 464 471

...

1 match found in sequence:
 US-09-252-991A-29231 ; Sequence 29231, Application US/09252991A
 (from "/srch/iaa/6B.COMB.pep")
 Sequence 29231, Application US/09252991A
 Patent No. 6551795

GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 29231
 LENGTH: 426

TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 Found using 'seq1' (new.key)

...

96 LPEATHLPGRKARKLPFRWRSAPFDGSEGRLEGAAAAPIVAVLARS
 116 123

...

1 match found in sequence:
 US-09-328-352-5472 ; Sequence 5472, Application US/09328352
 (from "/srch/iaa/6B.COMB.pep")
 Sequence 5472, Application US/09328352
 Patent No. 6562958

GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTE
 TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 5472
 LENGTH: 192

TYPE: PRT
 ORGANISM: Acinetobacter baumannii
 Found using 'seq1' (new.key)

...

65 TVLGVSSKQLQMDATGPLWKKQSFKGLAAGFTVSSLPGDKQSTL
 85 92

```
...
-----
1 match found in sequence:
US-09-663-600A-92 ; Sequence 92, Application US/09663600A
(from "/srch/iaa/6B.COMB.pep")
Sequence 92, Application US/09663600A
Patent No. 6573068
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean-Baptiste
APPLICANT: Duclert, Aymeric
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
FILE REFERENCE: 31.US3.CIP
CURRENT APPLICATION NUMBER: US/09/663,600A
CURRENT FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/191,997
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: 60/066,677
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/069,957
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/074,121
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/081,563
PRIOR FILING DATE: 1998-04-13
PRIOR APPLICATION NUMBER: 60/096,116
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/099,273
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 229
SOFTWARE: Patent.pm
SEQ ID NO 92
LENGTH: 230
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -24...-1
NAME/KEY: UNSURE
LOCATION: 54,79
OTHER INFORMATION: Xaa = any one of the twenty amino acids
Found using 'seq1' (new.key)
...
-----
10 GYILGLLGLTLVAMLLPSKTSYVGASIVTAVGFSKGLWMECATH
30 37
|-----|
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -24...-1
NAME/KEY: UNSURE
LOCATION: 54,79
OTHER INFORMATION: Xaa = any one of the twenty amino acids
Found using 'seq1' (new.key)
...
-----
1 match found in sequence:
US-09-663-600A-186 ; Sequence 186, Application US/09663600A
(from "/srch/iaa/6B.COMB.pep")
Sequence 186, Application US/09663600A
Patent No. 6573068
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, Jean-Baptiste
APPLICANT: Duclert, Aymeric
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
FILE REFERENCE: 31.US3.CIP
CURRENT APPLICATION NUMBER: US/09/663,600A
CURRENT FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/191,997
PRIOR FILING DATE: 1998-11-13
PRIOR APPLICATION NUMBER: 60/066,677
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/069,957
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/074,121
```

```
...
-----
1 match found in sequence:
US-09-732-210-332 ; Sequence 332, Application US/09732210
(from "/srch/iaa/6B.COMB.pep")
Sequence 332, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Mittanck, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 332
LENGTH: 70
TYPE: PRT
ORGANISM: Methanococcus jannaschii
Found using 'seq1' (new.key)
...
-----
1 MEWETCFCGCIETEPGKGMVVEKDTVLY
4 11
|-----|
1 match found in sequence:
US-09-107-532A-7233 ; Sequence 7233, Application US/09107532A
(from "/srch/iaa/6B.COMB.pep")
Sequence 7233, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
```

CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 7233:
SEQUENCE CHARACTERISTICS:
LENGTH: 354 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...354
SEQUENCE DESCRIPTION: SEQ ID NO: 7233:
Found using 'seq1' (new.key)

1 IIKLGRKMKLKIAPTGVSLLLGTLAACGSGSKDQ
12 19

1 match found in sequence:
US-09-107-532A-7307 ; Sequence 7307, Application US/09107532A
(from "/srch/iaa/6B.COMB.pep")
Sequence 7307, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 7307:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...68
SEQUENCE DESCRIPTION: SEQ ID NO: 7307:
Found using 'seq1' (new.key)

36 DRKLIIVLKSTHNNRNKWKYYSIGIRIS
56 63

1 match found in sequence:
US-09-215-418-2 ; Sequence 2, Application US/09215418
(from "/srch/iaa/6B.COMB.pep")
Sequence 2, Application US/09215418
Patent No. 6586217
GENERAL INFORMATION:
APPLICANT: Guimaraes, M. Jorge
APPLICANT: Bazan, J. Fernando
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: MAMMALIAN SELENOPHOSPHATE SYNTHETASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/215,418
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/406,359
FILING DATE: 17-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0493
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-9196
TELEFAX: 415-496-1200
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 amino acids
TYPE: amino acid

```

      TOPOLOGY: linear
      MOLECULE TYPE: protein
      Found using 'seq1' (new.key)
...
32      GRFSNVRPFPOTLGFSPSWRLTFSFGMKGCCKVPQETLLKLEGL
      {-----}
      52      59
...

-- Search Statistics --
Times:      CPU      Total Elapsed
           00:01:31.11      00:01:35.00

Number of sequences searched:      328807
Number of sequence hits:      36
Number of separate matches:      36
Number of sequence hits saved:      0
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:31:19 ; Search time 16 seconds

(without alignments)
48.084 Million cell updates/sec

Title: US-09-185-908-1

Perfect score: 23

Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	73.9	140	2	A72511
2	17	73.9	194	2	F72782
3	17	73.9	256	2	E64729
4	17	73.9	275	2	S70895
5	17	73.9	275	2	H82106
6	17	73.9	276	2	D87323
7	17	73.9	280	2	H83623
8	17	73.9	286	1	XYECCR
9	17	73.9	286	2	B90953
10	17	73.9	286	2	F95801
11	17	73.9	288	1	XYE3GM
12	17	73.9	288	2	AB0746
13	17	73.9	293	2	AG0204
14	17	73.9	293	2	B87303
15	17	73.9	308	2	B88392
16	17	73.9	379	2	T40384
17	17	73.9	531	2	S50965
18	17	73.9	593	2	A45191
19	17	73.9	1469	2	T09219
20	16	69.6	78	2	AC2678
21	16	69.6	78	2	H97459
22	16	69.6	91	2	JE0321
23	16	69.6	102	2	B34153
24	16	69.6	105	2	JC5414
25	16	69.6	129	2	S78153
26	16	69.6	130	2	S73215
27	16	69.6	130	2	S77479
28	16	69.6	130	2	S78255
29	16	69.6	130	2	T06945

30 16 69.6 131 2 A12329 30S ribosomal prot
31 16 69.6 131 2 E51276 hypothetical prote
32 16 69.6 133 2 AF3272 hypothetical prote
33 16 69.6 135 2 G87686 hypothetical prote
34 16 69.6 136 2 C95291 ribosomal protein
35 16 69.6 138 1 R3NT11 hypothetical prote
36 16 69.6 144 2 D72517 hypothetical prote
37 16 69.6 165 2 S35195 lipid A-myristate
38 16 69.6 173 2 AB3310 gene G protein - p
39 16 69.6 177 1 ZGBCP4 beta-lytic metallo
40 16 69.6 178 1 LYXB4 hypothetical prote
41 16 69.6 179 2 H72697 IS 426 transposase
42 16 69.6 184 2 AB3163 conserved hypothet
43 16 69.6 199 2 AB3195 conserved hypothet
44 16 69.6 201 2 G95852 hypothetical prote
45 16 69.6 201 2 F95270

ALIGNMENTS

RESULT 1

A72511
hypothetical protein APE2063 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: A72511
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tanaka, T.; Tanaka, T.; Fuchashii, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Ae
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: A72511
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-140 <KAW>
A:Cross-references: DBJ:AF000063; NID:95105654; PIDN:BA81073.1; PID:dl044859; PID
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2063

Query Match 73.9%; Score 17; DB 2; Length 140;
Best Local Similarity 25.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
Db 79 WSSASSSG 86

RESULT 2

F72782
hypothetical protein APE0247 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: F72782
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tanaka, T.; Tanaka, T.; Fuchashii, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72782
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-194 <KAW>
A:Cross-references: DBJ:AF000058; NID:95103388; PIDN:BA79160.1; PID:dl042936; PID:9
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0247

Query Match 73.9%; Score 17; DB 2; Length 194;
Best Local Similarity 25.0%; Pred. No. 1.3e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 15 WSAATASG 22

RESULT 3

E84729

Chemotaxis protein methyltransferase - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: E84729
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84729
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-256 <STO>
A:Cross-references: GB:AE002093; NID:g4263704; PIDN:AAZ15390.1; GSPDB:GN00139
C:Genetics:
A:Gene: Atg32150
A:Map position: 2

Query Match 73.9%; Score 17; DB 2; Length 256;
Best Local Similarity 25.0%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 222 WATATATG 229

RESULT 4

S70895

Chemotactic methyltransferase - Vibrio anguillarum
C:Species: Vibrio anguillarum
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 18-Jun-1999
C:Accession: S70895
R:O'Toole, R.; Milton, D.L.; Wolf-Watz, H.
Mol. Microbiol. 19, 625-637, 1996
A:Title: Chemotactic motility is required for invasion of the host by the fish pathogen
A:Reference number: S70894; MUID:96228710; PMID:8830252
A:Accession: S70895
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-275 <OTO>
A:Cross-references: GB:U36378; EMBL:L47344; NID:g1020321; PIDN:AAB38489.1; PID:g1177141
C:Genetics:
A:Gene: cher
C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransferase
F:4-261/Domain: protein-glutamate O-methyltransferase homology <PGM>

Query Match 73.9%; Score 17; DB 2; Length 275;
Best Local Similarity 25.0%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 108 WSAASSG 115

RESULT 5

H82106

Chemotaxis protein methyltransferase Cher VC2201 [imported] - Vibrio cholerae (strain N1
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: H82106
R:Heidelberg, V.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: H82106
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-275 <HEI>
A:Cross-references: GB:AE004292; GB:AE003852; NID:g9656760; PIDN:AAF95346.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N15961; biotype El Tor
C:Genetics:
A:Gene: VC2201
A:Map position: 1
C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltrans

Query Match 73.9%; Score 17; DB 2; Length 275;
Best Local Similarity 25.0%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 108 WSAASSG 115

RESULT 6

D87323

Chemotaxis protein methyltransferase Cher [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: D87323
R:Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; K
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87323
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-276 <STO>
A:Cross-references: GB:AE005673; NID:g13421800; PIDN:AAK22584.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0598
C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltr

Query Match 73.9%; Score 17; DB 2; Length 276;
Best Local Similarity 25.0%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 103 WSAATASG 110

RESULT 7

H83623

Probable chemotaxis protein methyltransferase PA0175 [imported] - Pseudomonas aerug
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83623
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83623
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-280 <STO>
A:Cross-references: GB:AE004455; GB:AE004091; NID:g9946002; PIDN:AAG03565.1; GSPDB:
A:Experimental source: strain PA01
C:Genetics:

A:Gene: PAW75
C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransferase

Query Match 73.9%; Score 17; DB 2; Length 280;
Best Local Similarity 25.0%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 116 WSAASTG 123

RESULT 8
XYECCR
protein-glutamate O-methyltransferase (EC 2.1.1.80) - Escherichia coli (strain K-12)
N:Alternate names: methyl-accepting chemotaxis protein O-methyltransferase; protein methyltransferase; Escherichia coli
C:Date: 28-Dec-1987 #sequence_revision 31-Oct-1997 #text_change 01-Mar-2002
C:Accession: D64951; C25195
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D64951
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-286 <BLAT>
A:Cross-references: GS:AE000282; GB:U00096; NID:g1788189; PIDN:AC74954.1; PID:g1788189;
A:Experimental source: strain K-12, substrain MG1655
R:Mutoh, N.; Simon, M.I.
J. Bacteriol. 165, 161-166, 1986
A:Title: Nucleotide sequence corresponding to five chemotaxis genes in Escherichia coli.
A:Reference number: A91811; MUID:86085665; PMID:3510184
A:Accession: C25195
A:Molecule type: DNA
A:Residues: 1-112, 'G', 114-286 <MUT>
A:Cross-references: GB:M13463; NID:g145517; PIDN:AAA23568.1; PID:g145523
C:Comment: This enzyme catalyzes the transfer of methyl groups from S-adenosylmethionine in MCP. The MCP methylation state of the cell is crucial for sensory responses and adaptation.
C:Genetics: chet
A:Gene: chet
A:Map position: 42 min
C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransferase
C:Keywords: chemotaxis; methyltransferase; S-adenosylmethionine
F:20-273/Domain: protein-glutamate O-methyltransferase homology <PGM>

Query Match 73.9%; Score 17; DB 1; Length 286;
Best Local Similarity 25.0%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 120 WSAASTG 127

RESULT 9
B90953
response regulator for chemotaxis [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: B90953
R:Hayaishi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B90953
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <RAY>
A:Cross-references: GB:BA000007; PIDN:BA836017.1; PID:g13362062; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIND 0509952

C:Genetics: chet
A:Gene: RCS2594
C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransferase

Query Match 73.9%; Score 17; DB 2; Length 286;
Best Local Similarity 25.0%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 120 WSAASTG 127

RESULT 10
F83801
response regulator for chemotaxis [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
C:Accession: F83801
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Maillier, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F83801
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 <STO>
A:Cross-references: GS:AE005174; NID:g12515946; PIDN:AA556874.1; GSPDB:GN00145; UWGP
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics: chet
A:Gene: chet
C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransferase

Query Match 73.9%; Score 17; DB 2; Length 286;
Best Local Similarity 25.0%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 120 WSAASTG 127

RESULT 11
XYEBGM
protein-glutamate O-methyltransferase (EC 2.1.1.80) - Salmonella typhimurium
N:Alternate names: methyl-accepting chemotaxis protein O-methyltransferase; protein methyltransferase; Salmonella typhimurium
C:Species: Salmonella typhimurium
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 05-May-2000
C:Accession: A29303
R:Simms, S.A.; Stock, A.M.; Stock, J.B.
J. Biol. Chem. 262, 8537-8543, 1987
A:Title: Purification and characterization of the S-adenosylmethionine: glutamyl met
A:Reference number: A29303; MUID:87250466; PMID:3298235
A:Accession: A29303
A:Molecule type: DNA
A:Residues: 1-288 <SIM>
A:Cross-references: GB:J02757; NID:g153902; PIDN:AAA27035.1; PID:g153903
C:Comment: This enzyme catalyzes the transfer of methyl groups from S-adenosylmethionine in MCP. The MCP methylation state of the cell is crucial for sensory responses and adaptation.
C:Genetics: chet
A:Gene: chet
A:Map position: 40 min
C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransferase
C:Keywords: chemotaxis; methyltransferase; S-adenosylmethionine; sensory transduction
F:20-273/Domain: protein-glutamate O-methyltransferase homology <PGM>

Query Match 73.9%; Score 17; DB 1; Length 288;
Best Local Similarity 25.0%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 120 WSAASTG 127

Db 120 WSAASTG 127

RESULT 12

AB0746

protein-glutamate O-methyltransferase (EC 2.1.1.80) - Salmonella enterica subsp. enteric
 C:Species: Salmonella enterica subsp. enterica serovar typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AB0746

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 the T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gea, P.
 Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AB0746

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-288 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD05669.1; PID:gi6503164; GSPDB:GNO0176

C:Genetics:

C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransferase
 C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 73.9%; Score 17; DB 2; Length 288;
 Best Local Similarity 25.0%; Pred. No. 1.7e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 120 WSAASTG 127

RESULT 13

AG0204

protein-glutamate O-methyltransferase (EC 2.1.1.80) [imported] - Yersinia pestis (strain
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 11-Jan-2002
 C:Accession: AG0204

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchall, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 ill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AG0204

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-290 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC90499.1; PID:gi5979710; GSPDB:GNO0175

C:Genetics:

C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltransferase
 C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 73.9%; Score 17; DB 2; Length 290;
 Best Local Similarity 25.0%; Pred. No. 1.8e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 123 WSAASTG 130

RESULT 14

B87303

chemotaxis protein methyltransferase CheR [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C:Accession: B87303

R.Nierman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: AB7249; MUID:21173698; PMID:11259647

A:Accession: B87303

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-293 <STO>

A:Cross-references: GB:AB005673; NID:gl3421602; PIDN:AAK2422.1; GSPDB:GNO0148

C:Genetics:

C:Superfamily: protein-glutamate O-methyltransferase; protein-glutamate O-methyltrans

Query Match 73.9%; Score 17; DB 2; Length 293;
 Best Local Similarity 25.0%; Pred. No. 1.8e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 127 WSAASSG 134

RESULT 15

B88392

protein R06B10.3 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: B88392

R:anonymous, the C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
 A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_

A:Notes: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;

A:Accession: B88392

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <STO>

A:Cross-references: GB:chr_III; PIDN:AAB95036.1; PID:g2746882; GSPDB:GNO00021; CESP:RO

C:Genetics:

A:Gens: R06B10.3

A:Map position: 3

Query Match 73.9%; Score 17; DB 2; Length 308;
 Best Local Similarity 25.0%; Pred. No. 1.8e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 115 WASSSTAG 122

Search completed: August 16, 2003, 14:31:45
 Job time : 18 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: August 16, 2003, 14:31:20 ; Search time 23 Seconds
(without alignments)
16.357 Million cell updates/sec

Title: US-09-185-908-1
Perfect score: 23
Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17	73.9	275	1	CHER_VIBAN
2	17	73.9	275	1	CHER_VIBPA
3	17	73.9	275	1	CHRL_VIBCH
4	17	73.9	280	1	CHRT_PSEAE
5	17	73.9	286	1	CHER_ECOLI
6	17	73.9	288	1	CHER_SALTY
7	17	73.9	379	1	AMPL_SCHPO
8	17	73.9	393	1	PTRR_HUMAN
9	16	69.6	105	1	CORT_HUMAN
10	16	69.6	130	1	RR11_CYPAP
11	16	69.6	130	1	RR11_ODOSI
12	16	69.6	130	1	RR11_PORPU
13	16	69.6	130	1	RS11_SYNEL
14	16	69.6	130	1	RS11_SYNP6
15	16	69.6	130	1	RS11_SYNP3
16	16	69.6	131	1	RS11_ANASP
17	16	69.6	138	1	RR11_TOBAC
18	16	69.6	177	1	VGG_BPG4
19	16	69.6	178	1	PRLE_LYSEN
20	16	69.6	210	1	YAC9_MAIZE
21	16	69.6	294	1	RLSA_SCHPO
22	16	69.6	294	1	RLSB_SCHPO
23	16	69.6	297	1	RL5_YEAST
24	16	69.6	301	1	RL5_NEUCR
25	16	69.6	310	1	LDH_BACHD
26	16	69.6	327	1	XYNC_EWEMI
27	16	69.6	328	1	YQ51_CAEEL
28	16	69.6	357	1	CADH_POPDE
29	16	69.6	374	1	PRLB_ACHLY
30	16	69.6	383	1	AMC1_ORYSA
31	16	69.6	385	1	OM7A_DROME
32	16	69.6	400	1	GUNS_BACAG
33	16	69.6	401	1	FLGE_ECOLI

34 16 69.6 402 1 FLGE_SALTY
35 16 69.6 405 1 FLGE_BUCAI
36 16 69.6 413 1 PEPS_STRTR
37 16 69.6 433 1 ELT2_CABEL
38 16 69.6 460 1 NU4M_CROLA
39 16 69.6 473 1 YV33_MYCLE
40 16 69.6 473 1 YV33_MYCTU
41 16 69.6 480 1 HEM3_EUGER
42 16 69.6 488 1 GUN1_BACS4
43 16 69.6 513 1 VNN1_HUMAN
44 16 69.6 527 1 RHGB_ASPAC
45 16 69.6 636 1 MAOC_MAIZE

P16322 salmonella
P57422 buchnera sp
Q9x4a7 streptococ
Q10655 caenorhabdi
P34194 crossostoma
P37391 mycobacteri
O06250 mycobacteri
P13416 euglena gra
P06566 bacillus sp
O95497 homo sapien
Q00019 aspergillus
P16243 zea mays (m

ALIGNMENTS

RESULT 1
CHER_VIBAN
ID CHER_VIBAN STANDARD; PRT; 275 AA.
AC Q57508;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chemotaxis protein methyltransferase (EC 2.1.1.80).
GN CHER
OS Vibrio anguillarum (Listonella anguillarum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Listonella.
OX NCBI_TaxID=55601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NB10 / Serotype O1;
RX MEDLINE=96228710; PubMed=8930252;
RA O'Toole R., Milton D.B., Wolf-Watz H.;
RT "Chemotactic motility is required for invasion of the host by the fish
pathogen Vibrio anguillarum."
RL Mol. Microbiol. 19:625-637(1996).
CC -!- FUNCTION: METHYLATION OF THE MEMBRANE-BOUND METHYL-ACCEPTING
CHEMOTAXIS PROTEINS (MCP) TO FORM GAMMA-GLUTAMYL METHYL ESTER
RESIDUES IN MCP (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-glutamate
= S-adenosyl-L-homocysteine + protein L-glutamate methyl ester.
CC -!- SIMILARITY: Contains 1 cher-type methyltransferase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U36378; AAC38489.1; -;
CC PIR: S70895; S70895.
CC HSSP: P07801; 1AF7.
CC InterPro: IPR000780; Cher Metranf.
CC InterPro: IPR001601; Methyltransf.
CC Pfam: PF01739; Cher; 1.
CC Pfam: PF03705; Cher_N; 1.
CC PRINTS: PR00996; CHERMTFRASE.
CC SMART: SM00138; Metrc; 1.
CC PROSITE: PS0123; CHER; 1.
KW Transferase; Methyltransferase.
FT DOMAIN 1 275 CHER-TYPE METHYLTRANSFERASE.
SQ SEQUENCE 275 AA; 30808 MW; 0A44C10F0C5D9C53 CRC64;

Query Match 73.9%; Score 17; DB 1; Length 275;
Best Local Similarity 25.0%; Pred. No. 8.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|

Db 108 WSAASSG 115

RESULT 2

CHER_VIBCHA
ID CHER_VIBCHA STANDARD; PRT: 275 AA.
AC Q9X9K2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Chemotaxis protein methyltransferase (EC 2.1.1.80).
GN CHER OR VP0774.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BB22;
RA McCarter L.L.;
RT "Polar flagellar region I";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae";
RL Lancet 361:743-749(2003).
CC -!- FUNCTION: METHYLATION OF THE MEMBRANE-BOUND METHYL-ACCEPTING
CC CHROMATIN PROTEINS (MCP) TO FORM GAMMA-GLUTAMYL METHYL ESTER
CC RESIDUES IN MCP (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-glutamate
CC = S-adenosyl-L-homocysteine + protein L-glutamate methyl ester.
CC -!- SIMILARITY: Contains 1 cheR-type methyltransferase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U12817; AAD42911.1; -;
CC EMBL; AP005075; BAC59037.1; -;
CC WSSP; P07801; IAF7
CC InterPro; IPR000780; ChR_Metranf.
CC InterPro; IPR001601; Methyltransf.
CC Pfam; PF01739; ChR; 1.
CC Pfam; PF03705; ChR_N; 1.
CC PRINTS; PR00996; ChR_MTFRASE.
CC SMART; SM00138; Metrc; 1.
CC PROSITE; PS50123; ChR; 1.
KW Transferrase; Methyltransferase; Complete proteome.
FT DOMAIN 1 275 CHER-TYPE METHYLTRANSFERASE.
SQ SEQUENCE 275 AA; 30822 MW; 85102B53E5D53794 CRC64;

Query Match 73.9%; Score 17; DB 1; Length 275;
Best Local Similarity 25.0%; Pred. No. 8.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 108 WSAASSG 115

RESULT 3

CHRL_VIBCH
ID CHRL_VIBCH STANDARD; PRT: 275 AA.

Q9KQ06; Q9XCL5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chemotaxis protein methyltransferase 1 (EC 2.1.1.80).
GN CHER1 OR CHER OR VC2201.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisele J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Fetscherin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae";
RL Nature 406:477-483(2000).
RN [2]
RP SEQUENCE OF 10-266 FROM N.A.
RC STRAIN=CVD110;
RX MEDLINE=99328977; PubMed=10400589;
RA O'Toole R., Lundberg S., Fredriksson S.A., Jansson A., Nilsson B.,
RA Wolf-Watz H.;
RT "The chemotactic response of Vibrio anguillarum to fish intestinal
mucus is mediated by a combination of multiple mucus components";
RL J. Bacteriol. 181:4308-4317(1999).
CC -!- FUNCTION: METHYLATION OF THE MEMBRANE-BOUND METHYL-ACCEPTING
CC CHROMATIN PROTEINS (MCP) TO FORM GAMMA-GLUTAMYL METHYL ESTER
CC RESIDUES IN MCP (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-glutamate
CC = S-adenosyl-L-homocysteine + protein L-glutamate methyl ester.
CC -!- SIMILARITY: Contains 1 cheR-type methyltransferase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE004291; AAF95346.1; -;
CC EMBL; AF139167; AAD45254.1; -;
CC FIR; H82106; H82106.
CC HSSP; P07801; IAF7.
CC TIGR; VC2201; -;
CC InterPro; IPR000780; ChR_Metranf.
CC InterPro; IPR001601; Methyltransf.
CC Pfam; PF01739; ChR; 1.
CC Pfam; PF03705; ChR_N; 1.
CC PRINTS; PR00996; ChR_MTFRASE.
CC SMART; SM00138; Metrc; 1.
CC PROSITE; PS50123; ChR; 1.
KW Transferrase; Methyltransferase; Complete proteome.
FT DOMAIN 1 275 CHER-TYPE METHYLTRANSFERASE.
SQ SEQUENCE 275 AA; 30871 MW; F008ADCBFA46A921 CRC64;

Query Match 73.9%; Score 17; DB 1; Length 275;
Best Local Similarity 25.0%; Pred. No. 8.5e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 108 WSAASSG 115

RESULT 4

```

CH2_PFEAE
ID CHR2_PSEAE STANDARD; PRT; 280 AA.
AC Q916V7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chemotaxis protein methyltransferase 2 (EC 2.1.1.80).
GN CH22 OR PA0175.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Phan X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey R.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.B., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Relzer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -1- FUNCTION: METHYLATION OF THE MEMBRANE-BOUND METHYL-ACCEPTING
CC CHEMOTAXIS PROTEINS (MCP) TO FORM GAMMA-GLUTAMYL METHYL ESTER
CC RESIDUES IN MCP (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-glutamate
CC = S-adenosyl-L-homocysteine + protein L-glutamate methyl ester.
CC -1- SIMILARITY: Contains 1 cher-type methyltransferase domain.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE004455; AG03565.1;
CC PIR; H83623; H83623.
CC HSSP; P07801; IAF7.
CC InterPro; IPR000780; Cher_Metranf.
CC InterPro; IPR001601; Methyltransf.
CC Pfam; PF01739; Cher; 1.
CC Pfam; PF03705; Cher_N; 1.
CC PRINTS; PR00996; CHERMTFRASE.
CC SMART; SM00138; Metrc; 1.
CC PROSITE; PS0123; Cher; 1.
CC TRANSFERASE; Methyltransferase; Complete proteome.
FT DOMAIN 10 280 CHER-TYPE METHYLTRANSFERASE.
SQ SEQUENCE 280 AA; 32028 MW; FD2AF1491DDC57FE CRC64;

Query Match 73.9%; Score 17; DB 1; Length 280;
Best Local Similarity 25.0%; Pred. No. 8.6e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
Db 116 WSAASTG 123

RESULT 5
CH2_ECOLI
ID CH2_ECOLI STANDARD; PRT; 286 AA.
AC F07364; F78071;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chemotaxis protein methyltransferase (EC 2.1.1.80).
GN CH2 OR CHEX OR B1884.

```

```

OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86085665; PubMed=3510184;
RA Watch N., Simon M.L.;
RT "Nucleotide sequence corresponding to five chemotaxis genes in
RT Escherichia coli."
RL J. Bacteriol. 165:161-166(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
RA Saito N., Sampei G., Seki Y., Sivasubram S., Tagami H.,
RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map."
RL DNA Res. 3:379-392(1996).
CC -1- FUNCTION: METHYLATION OF THE MEMBRANE-BOUND METHYL-ACCEPTING
CC CHEMOTAXIS PROTEINS (MCP) TO FORM GAMMA-GLUTAMYL METHYL ESTER
CC RESIDUES IN MCP.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-glutamate
CC = S-adenosyl-L-homocysteine + protein L-glutamate methyl ester.
CC -1- MISCELLANEOUS: THE MCP METHYLATION STATE OF THE CELL IS CRUCIAL
CC FOR SENSORY RESPONSES AND ADAPTATIONS.
CC -1- SIMILARITY: Contains 1 cher-type methyltransferase domain.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M13463; AAA23568.1;
CC EMBL; AE000282; AAC74954.1;
CC EMBL; D90830; BAA15700.1;
CC EMBL; D90831; BAA15705.1;
CC PIR; D64951; AYECCR.
CC HSSP; P07801; IAF7.
CC EcoGene; EG10148; Cher.
CC InterPro; IPR000780; Cher_Metranf.
CC InterPro; IPR001601; Methyltransf.
CC InterPro; IPR000051; SAM_bind.
CC Pfam; PF01739; Cher; 1.
CC Pfam; PF03705; Cher_N; 1.
CC PRINTS; PR00996; CHERMTFRASE.
CC SMART; SM00138; Metrc; 1.
CC PROSITE; PS0123; Cher; 1.
CC TRANSFERASE; Methyltransferase; Complete proteome.
FT DOMAIN 15 286 CHER-TYPE METHYLTRANSFERASE.
SQ CONFLICT 113 113 R -> G (IN REF. 1).
SQ SEQUENCE 286 AA; 32849 MW; 2AFA307DD406B135 CRC64;

Query Match 73.9%; Score 17; DB 1; Length 286;
Best Local Similarity 25.0%; Pred. No. 8.7e+02;

```

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 120 WSAASTG 127

RESULT 6

CHER_SALTY STANDARD; PRT; 288 AA.

AC 207801; 1

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Chemotaxis protein methyltransferase (EC 2.1.1.80).

GN CHER OR STM1918

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI_TaxID=602;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87250466; PubMed=3298235;

RA Shims S.A., Stock A.M., Stock J.B.;

RT "Purification and characterization of the

RT S-adenosylmethionine:glutaryl methyltransferase that modifies

RT membrane chemoreceptor proteins in bacteria.;"

RL J. Biol. Chem. 262:8537-8543(1987).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-LT2 / SGSC1412 / ATCC 700720;

RA MEDLINE=21534948; PubMed=11677609;

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,

RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,

RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,

RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,

RA Waterston R., Wilson R.K.;

RT "Complete genome sequence of Salmonella enterica serovar Typhimurium

RT LT2.;"

RL Nature 413:852-856(2001).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 11-284.

RX MEDLINE=97277239; PubMed=9115443;

RA Djordjevic S., Stock A.M.;

RT "Crystal structure of the chemotaxis receptor methyltransferase ChEr

RT suggests a conserved structural motif for binding

RT S-adenosylmethionine.;"

RL Structure 5:545-558(1997).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 16-284.

RX MEDLINE=98290446; PubMed=9628482;

RA Djordjevic S., Stock A.M.;

RT "Chemotaxis receptor recognition by protein methyltransferase ChEr.;"

RL Nat. Struct. Biol. 5:446-450(1998).

CC -1- FUNCTION: METHYLATION OF THE MEMBRANE-BOUND METHYL-ACCEPTING

CC CHEMOTAXIS PROTEINS (MCP) TO FORM GAMMA-GLUTAMYL METHYL ESTER

CC RESIDUES IN MCP.

CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + protein L-glutamate

CC = S-adenosyl-L-homocysteine + protein L-glutamate methyl ester.

CC -1- MISCELLANEOUS: THE MCP METHYLATION STATE OF THE CELL IS CRUCIAL

CC FOR SENSORY RESPONSES AND ADAPTATIONS.

CC -1- SIMILARITY: Contains 1 chEr-type methyltransferase domain.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; J02757; AAA27035.1; -

DR EMBL; AB008785; AAL20834.1; -

DR PIR; A29303; XVEBGM.

DR PDB; 1AF7; 28-JAN-98.

DR PDB; 1BC5; 13-JAN-99.

DR StyGene; SGI0054; cher.

DR InterPro; IPR000780; Cher_Metranf.

DR InterPro; IPR001601; Methyltransf.

DR InterPro; IPR000051; SAM_bind.

DR Pfam; PF01739; Cher; 1.

DR Pfam; PF03705; Cher_N; 1.

DR PRINTS; PR00996; CHERMTFRASE.

DR SMART; SM00138; Metrc; 1.

DR PROSITE; PS00123; CHER; 1.

KW Transferase; Methyltransferase; Chemotaxis; 3D-structure;

KW Complete proteome.

FT DOMAIN 15 286

FT HELIX 24 38

FT HELIX 44 46

FT HELIX 47 61

FT TURN 62 62

FT HELIX 66 75

FT TURN 77 78

FT TURN 80 80

FT HELIX 81 89

FT TURN 96 101

FT HELIX 102 112

FT STRAND 117 121

FT TURN 125 127

FT HELIX 128 141

FT TURN 145 146

FT STRAND 148 154

FT TURN 157 165

FT TURN 166 166

FT STRAND 167 169

FT HELIX 170 173

FT TURN 174 175

FT HELIX 178 184

FT STRAND 185 187

FT TURN 190 191

FT STRAND 195 198

FT HELIX 200 203

FT TURN 204 205

FT STRAND 206 210

FT TURN 213 214

FT STRAND 224 229

FT HELIX 233 235

FT HELIX 238 248

FT HELIX 249 251

FT STRAND 252 260

FT TURN 262 263

FT TURN 267 269

FT TURN 271 272

FT STRAND 273 277

FT TURN 278 279

FT STRAND 280 283

SQ SEQUENCE 288 AA; 32923 MW; 4D344E6F326DD482 CRC64;

Query Match 73.9%; Score 17; DB 1; Length 288;

Best Local Similarity 25.0%; Pred. No. 8.8e+02;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 120 WSAASTG 127

RESULT 7

AMPI_SCHPO STANDARD; PRT; 379 AA.

ID AMPI_SCHPO

AC 059730;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable methionine aminopeptidase 1 (EC 3.4.11.18) (MetAP 1)


```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=97236300; PubMed=9125122;
RX Fukusumi S., Kitada C., Takekawa S., Kizawa H., Sakamoto J.,
RA Miyamoto M., Hinuma S., Kitano K., Fujino M.;
RT "Identification and characterization of a novel human cortistatin-like
RT peptide.";
RL Biochem. Biophys. Res. Commun. 232:157-163(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97349120; PubMed=9205124;
RA de Lecea L., Ruiz-Lozano P., Danielson P.E., Peelle-Kirley J.,
RA Poye P.E., Frankel W.N., Sutcliffe J.G.;
RT "Cloning, mRNA expression, and chromosomal mapping of mouse and human
RT precortistatin.";
RL Genomics 42:499-506(1997).
CC -!- FUNCTION: BINDS TO ALL HUMAN SOMATOSTATIN RECEPTOR (SSTR)
CC SUBTYPES. IT ALSO INHIBITS CAMP PRODUCTION INDUCED BY FORSKOLIN
CC THROUGH SSTRS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A SUBSET OF GABAERGIC CELLS IN
CC THE CORTEX AND HIPPOCAMPUS.
CC -!- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AB000263; BAA19770.1; -
DR EMBL: AF013252; AAB66895.1; -
DR FIRM: J05414; J05414.
DR GENE: HGNC:2257; CORT.
DR MIM: 602784; -
DR GO: GO:0005625; C:soluble fraction; TAS.
DR GO: GO:0005184; F:neuropeptide hormone activity; TAS.
DR GO: GO:0007193; P:G-protein signaling, adenylate cyclase inh. . . ; TAS.
DR GO: GO:0007268; P:synaptic transmission; TAS.
DR InterPro: IPR004250; Somatostatin; 1.
DR Pfam: PF03002; Somatostatin; 1.
DR Cleavage on pair of basic residues; Hormone; Signal.
KW SIGNAL 1 18 POTENTIAL.
FT PEPTIDE 77 105 CORTISTATIN-29 (POTENTIAL).
FT PEPTIDE 89 105 CORTISTATIN-17.
FT DISULFID 93 104 BY SIMILARITY.
SQ SEQUENCE 105 AA; 11532 MW; 09578F4520201551 CRC64;

Query Match 69.6%; Score 16; DB 1; Length 105;
Best Local Similarity 25.0%; Pred. No. 7.7e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 56 WTSQASAG 63

RESULT 10
RR11_CVAPA STANDARD; PRT; 130 AA.
AC P48136;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cyanelle 30S ribosomal protein S11.
GN RPS11.
OS Cyanophora paradoxa.
OG Cyanelle.
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.

```

```

OX NCBI_TaxID=2762;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UTEX LB 555 / Pringsheim;
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RA Bryant D.A.;
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";
RL Plant Mol. Biol. Rep. 13:327-332(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=UTEX LB 555 / Pringsheim;
RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
RA Farley J.V., Schluchter W.M., Chung S., Newmann-Spallart C.,
RA Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
RT the genetic complexity of a primitive plastid.";
RL (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL Schwemmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
RL (1997).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit.
CC -!- SUBCELLULAR LOCATION: Cyanelle.
CC -!- SIMILARITY: BELONGS TO THE SLIP FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U30821; AAA81288.1; -
DR PIR: T06945; T06945.
DR HAMAP: MF_01310; -; 1.
DR InterPro: IPR001971; Ribosomal_S11.
DR Pfam: PF00411; Ribosomal_S11; 1.
DR ProDom: PD001010; Ribosomal_S11; 1.
DR PROSITE: PS00054; RIBOSOMAL_S11; 1.
DR Ribosome: R00000; RNA-binding; rRNA-binding; Cyanelle.
SQ SEQUENCE 130 AA; 13849 MW; B544027CF4D2AC09 CRC64;

Query Match 69.6%; Score 16; DB 1; Length 130;
Best Local Similarity 25.0%; Pred. No. 9.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 45 WASAGSSG 52

RESULT 11
RR11_ODOSI STANDARD; PRT; 130 AA.
AC P49499;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S11.
GN RPS11.
OS Odontella sinensis (Marine centric diatom).
OG Chloroplast.
OC Eukaryota; Stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Bidulphiophycidae; Eupodiscaceae; Eupodiscaceae; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RP SEQUENCE FROM N.A.
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis.";
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit.
CC -!- SUBCELLULAR LOCATION: Chloroplast.

```



```
CC -!- SIMILARITY: BELONGS TO THE SLIP FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 267753; CAA91628.1; -
DR FIRM; S78255; S78255.
DR HAMAP; MF_01310; -; 1.
DR InterPro; IPR001971; Ribosomal_S11.
DR Pfam; PF00411; Ribosomal_S11; 1.
DR ProDom; PD001010; Ribosomal_S11; 1.
DR PROSITE; PS00054; RIBOSOMAL_S11; 1.
DR Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
KW RIBOSOMAL protein; RNA-binding; rRNA-binding; 8FAP02371B1565CA CRC64;
SQ SEQUENCE 130 AA; 13772 MW; 8FAP02371B1565CA CRC64;
SQ -----
Query Match 69.6%; Score 16; DB 1; Length 130;
Best Local Similarity 25.0%; Pred. No. 9.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 WXXXXXXG 8
DB 45 WSSGSG 52
-----
RESULT 12
RR11_PORPU STANDARD; PRT; 130 AA.
AC P51294;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S11.
GN RPS11.
OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RT Keith M.E., Munkholland J.;
RL "Complete nucleotide sequence of the Porphyra purpurea chloroplast
genome."
RT Plant Mol. Biol. Rep. 13:333-335(1995).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: BELONGS TO THE SLIP FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38804; AAC08180.1; -
DR FIRM; S73215; S73215.
DR HAMAP; MF_01310; -; 1.
DR InterPro; IPR001971; Ribosomal_S11.
DR Pfam; PF00411; Ribosomal_S11; 1.
DR ProDom; PD001010; Ribosomal_S11; 1.
DR PROSITE; PS00054; RIBOSOMAL_S11; 1.
DR Ribosomal protein; RNA-binding; rRNA-binding; Chloroplast.
KW RIBOSOMAL protein; RNA-binding; 7D9E4B3944149C73 CRC64;
SQ SEQUENCE 130 AA; 13803 MW; 7D9E4B3944149C73 CRC64;
SQ -----
Query Match 69.6%; Score 16; DB 1; Length 130;
Best Local Similarity 25.0%; Pred. No. 9.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 WXXXXXXG 8
DB 45 WSSGSG 52
-----
RESULT 13
RS11_SYNEL STANDARD; PRT; 130 AA.
AC P59379;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S11.
GN RPSK OR RPS11 OR TIR0104.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimpō S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:123-130(2002).
CC -!- FUNCTION: Located on the platform of the 30S subunit, it bridges
CC several disparate RNA helices of the 16S rRNA. Forms part of the
CC Shine-Dalgarno cleft in the 70S ribosome (By similarity).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Interacts with
CC proteins S7 and S18. Binds to IF-3 (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SLIP FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF005369; BAC07657.1; -
DR HAMAP; MF_01310; -; 1.
DR Pfam; PF00411; Ribosomal_S11; 1.
DR ProDom; PD001010; Ribosomal_S11; 1.
DR PROSITE; PS00054; RIBOSOMAL_S11; 1.
DR Ribosomal protein; RNA-binding; rRNA-binding; Complete proteome.
KW RIBOSOMAL protein; RNA-binding; 956D26EBE9417165 CRC64;
SQ SEQUENCE 130 AA; 13711 MW; 956D26EBE9417165 CRC64;
SQ -----
Query Match 69.6%; Score 16; DB 1; Length 130;
Best Local Similarity 25.0%; Pred. No. 9.1e+02;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 WXXXXXXG 8
DB 45 WASAGSG 52
-----
RESULT 14
RS11_SYNP6 STANDARD; PRT; 130 AA.
AC O24709;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S11.
GN RPSK OR RPS11.
OS Synechococcus sp. (strain PCC 6301) (Anacystis nidulans).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
```

```
OX NCBI_TaxID=1139;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97444291; PubMed=9300823;
RA Sugita M., Sugishita H., Fujishiro T., Tsuboi M., Sugita C., Endo T.,
RA Sugita M.;
RT "Organization of a large gene cluster encoding ribosomal proteins in
RT the cyanobacterium Synechococcus sp. strain PCC 6301: comparison of
RT gene clusters among cyanobacteria, eubacteria and chloroplast
RT genomes.";
RL Gene 195:73-79(1997).
CC -!- FUNCTION: Located on the platform of the 30S subunit, it bridges
CC several disparate RNA helices of the 16S rRNA. Forms part of the
CC Shine-Dalgarno cleft in the 70S ribosome (By similarity).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Interacts with
CC proteins S7 and S18. Binds to IF-3 (By similarity).
CC -!- SIMILARITY: BELONGS TO THE S11P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB000111; BAA22471.1; -
CC DR HAMAP; MF_01310; -; 1.
CC DR InterPro; IPR001971; Ribosomal_S11.
CC DR Pfam; PF00411; Ribosomal_S11; 1.
CC DR PRODOM; PD001010; Ribosomal_S11; 1.
CC DR PROSITE; PS00054; RIBOSOMAL_S11; 1.
CC DR Ribosomal protein; rRNA-binding; rRNA-binding; Complete proteome.
CC KW RIBOSOMAL_S11; 130 AA; 13712 MW; 4FA9622B3C19B0F CRC64;
CC SEQUENCE 130 AA; 13712 MW; 4FA9622B3C19B0F CRC64;
CC
CC Query Match 69.6%; Score 16; DB 1; Length 130;
CC Best Local Similarity 25.0%; Pred. No. 9.1e+02;
CC Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 1 WXXXXXXG 8
CC |
CC DB 45 WASAGSSG 52
CC
CC RESULT 15
CC RS11_SYNY3
CC ID RS11_SYNY3 STANDARD; PRT; 130 AA.
CC AC P73298;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE 30S ribosomal protein S11.
CC GN RPSK OR RPS11 OR SLL1817.
CC OS Synechocystis sp. (strain PCC 6803).
CC OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
CC OX NCBI_TaxID=1148;
CC RN [1]
CC SEQUENCE FROM N.A.
CC RX MEDLINE=97061201; PubMed=8905231;
CC RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
CC Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
CC Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
CC Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
CC Yamada M., Yasuda M., Tabata S.;
CC "Sequence analysis of the genome of the unicellular cyanobacterium
CC Synechocystis sp. strain PCC6803. II. Sequence determination of the
CC entire genome and assignment of potential protein-coding regions.";
CC DNA Res. 3:109-136(1996).
CC -!- FUNCTION: Located on the platform of the 30S subunit, it bridges
CC several disparate RNA helices of the 16S rRNA. Forms part of the
CC Shine-Dalgarno cleft in the 70S ribosome (By similarity).
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Interacts with
CC proteins S7 and S18. Binds to IF-3 (By similarity).
CC
```

```
CC -!- SIMILARITY: BELONGS TO THE S11P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90905; BAA17326.1; -
CC DR PIR; S77479; S77479.
CC DR HAMAP; MF_01310; -; 1.
CC DR InterPro; IPR001971; Ribosomal_S11.
CC DR Pfam; PF00411; Ribosomal_S11; 1.
CC DR PRODOM; PD001010; Ribosomal_S11; 1.
CC DR PROSITE; PS00054; RIBOSOMAL_S11; 1.
CC DR Ribosomal protein; rRNA-binding; rRNA-binding; Complete proteome.
CC KW RIBOSOMAL_S11; 130 AA; 13762 MW; B275DD0FB5F59CE CRC64;
CC SEQUENCE 130 AA; 13762 MW; B275DD0FB5F59CE CRC64;
CC
CC Query Match 69.6%; Score 16; DB 1; Length 130;
CC Best Local Similarity 25.0%; Pred. No. 9.1e+02;
CC Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 1 WXXXXXXG 8
CC |
CC DB 45 WASAGSSG 52
CC
CC Search completed: August 16, 2003, 14:35:28
CC Job time : 25 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:31:20 ; Search time 96 Seconds
(without alignments)
21.504 Million cell updates/sec

Title: US-09-185-908-1
Perfect score: 23
Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_prodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	73.9	98	Q8RVA6	Q8RVA6 oryza sativ
2	17	73.9	140	Q9YA76	Q9YA76 aeropyrum p
3	17	73.9	194	Q9YFK1	Q9YFK1 aeropyrum p
4	17	73.9	263	Q9SKY5	Q9SKY5 arabidopsis
5	17	73.9	270	Q80539	Q80539 homo sapien
6	17	73.9	270	Q9WUJ5	Q9WUJ5 mus musculu
7	17	73.9	274	Q8GLP9	Q8GLP9 aeromonas h
8	17	73.9	275	Q8GM67	Q8GM67 vibrio fisc
9	17	73.9	275	Q8PIM6	Q8PIM6 xanthomonas
10	17	73.9	275	Q8PTA7	Q8PTA7 xanthomonas
11	17	73.9	275	Q8DFI4	Q8DFI4 vibrio vuln
12	17	73.9	276	Q9AAJ9	Q9AAJ9 caulobacter
13	17	73.9	283	Q8F114	Q8F114 leptospira
14	17	73.9	286	Q8XCF8	Q8XCF8 escherichia
15	17	73.9	286	Q8PIK6	Q8PIK6 xanthomonas
16	17	73.9	286	Q8FGP4	Q8FGP4 escherichia

17	17	73.9	288	16	Q8Z5V1	Q8Z5V1 salmonella
18	17	73.9	290	2	Q9FAD9	Q9FAD9 enterobacte
19	17	73.9	290	16	Q8ZFW1	Q8ZFW1 yersinia pe
20	17	73.9	293	2	Q87716	Q87716 caulobacter
21	17	73.9	293	16	Q9AB01	Q9AB01 caulobacter
22	17	73.9	295	16	Q8DOP3	Q8DOP3 yersinia pe
23	17	73.9	299	16	Q8NSK6	Q8NSK6 corynebacte
24	17	73.9	329	10	Q8H8U0	Q8H8U0 oryza sativ
25	17	73.9	333	4	Q8N5V1	Q8N5V1 homo sapien
26	17	73.9	334	2	Q9AEZ6	Q9AEZ6 frankia sp.
27	17	73.9	362	5	Q44871	Q44871 caenorhabdi
28	17	73.9	371	11	Q921Q4	Q921Q4 mus musculu
29	17	73.9	402	10	Q8LEB9	Q8LEB9 arabidopsis
30	17	73.9	402	10	Q9LUZ6	Q9LUZ6 arabidopsis
31	17	73.9	413	13	Q9DGN4	Q9DGN4 xenopus lae
32	17	73.9	432	2	Q8GGR5	Q8GGR5 streptomyce
33	17	73.9	519	10	Q8S232	Q8S232 oryza sativ
34	17	73.9	527	16	Q8VKQ6	Q8VKQ6 mycobacteri
35	17	73.9	531	3	Q12235	Q12235 saccharomyc
36	17	73.9	570	3	Q9C0U9	Q9C0U9 schizosacch
37	17	73.9	595	6	Q9TU31	Q9TU31 canis famil
38	17	73.9	704	11	Q8CI96	Q8CI96 mus musculu
39	17	73.9	964	4	Q8NHB4	Q8NHB4 homo sapien
40	17	73.9	1012	10	Q9AV30	Q9AV30 oryza sativ
41	17	73.9	1469	4	Q9Y6P7	Q9Y6P7 homo sapien
42	17	73.9	1478	17	Q8TLB5	Q8TLB5 methanosarc
43	16	69.6	58	2	Q49260	Q49260 mycoplasma
44	16	69.6	70	8	Q8LXQ5	Q8LXQ5 pimephales
45	16	69.6	70	8	Q8LUS9	Q8LUS9 pimephales

ALIGNMENTS

RESULT 1

Q8RVA6 ID Q8RVA6 PRELIMINARY; PRT; 98 AA.
AC Q8RVA6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE P0684C02.4 protein (P0557A01.36 protein).
GN P0684C02.4 OR P0557A01.36.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:p0684C02."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:p0557A01."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003290; BAB9048.1; -
DR EMBL; AP003280; BAB99785.1; -
DR Gramene; Q8RVA6; -
SQ SEQUENCE 98 AA; 9330 MW; 1AFE4FEB49EC7274 CRC64;

Query Match 73.9%; Score 17; DB 10; Length 98;

Best Local Similarity 25.08; Pred. No. 3e+03;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 78 WAAARATG 85

RESULT 2

Q9YA76 PRELIMINARY; PRT; 140 AA.
 AC Q9YA76;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DE Hypothetical protein APE2063.
 GN APE2063.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 CC Desulfurococcales; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kavarabavasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1";
 RL DNA Res. 6:83-101(1999).
 DR EMBL; AP000063; SRA81073.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 140 AA; 14621 MW; 77D5DE847B13CB AE CRC64;

Query Match 73.9%; Score 17; DB 17; Length 140;
 Best Local Similarity 25.0%; Pred. No. 4 le+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 79 WSSASSG 86

RESULT 3

Q9YFK1 PRELIMINARY; PRT; 194 AA.
 AC Q9YFK1;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DE Hypothetical protein APE0247.
 GN APE0247.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 CC Desulfurococcales; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kavarabavasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1";
 RL DNA Res. 6:83-101(1999).
 DR EMBL; AP000058; BAA79160.1; -.
 DR InterPro; IPR002155; Thiolase.
 DR PROSITE; PS00098; THIOLEASE_1; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 194 AA; 20154 MW; C2EBDBD1A0F429BA CRC64;

Query Match 73.9%; Score 17; DB 17; Length 194;
 Best Local Similarity 25.0%; Pred. No. 5.5e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 15 WASATASG 22

RESULT 4

Q9SKY5 PRELIMINARY; PRT; 263 AA.
 AC Q9SKY5; Q93VP6;
 DT 01-MAY-2003 (TrEMBLrel. 13, Created)
 DT 01-JUN-2003 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative hydrolase (At2g32150/F22D22.10).
 GN AT2G32150.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyte; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chenk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Havashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Chen H., Cheuk R.,
 RA Chung M.K., Kim C., Lin J., Liu S.X., Pham P.K., Sakano H., Shinn P.,
 RA Yamada K., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Chenk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis ORF clones";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA EMBL; AC006223; AAD15390.2; -.
 DR EMBL; AY057535; AAL09775.1; -.
 DR EMBL; AF370598; AAK13917.1; -.
 DR EMBL; AY093978; AAM16239.1; -.
 DR InterPro; IPR006402; HAD-SF-IA-V3.
 DR InterPro; IPR005834; Hydrolase.
 DR Pfam; PF00702; Hydrolase; 1.
 DR TIGRFAMs; TIGR01509; HAD-SF-IA-V3; 1.
 KW Hydrolase.
 SQ SEQUENCE 263 AA; 29036 MW; 9684DD73006D6D22 CRC64;

Query Match 73.9%; Score 17; DB 10; Length 263;
 Best Local Similarity 25.0%; Pred. No. 7.2e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 229 WATATATG 236

RESULT 5

AC O60539 PRELIMINARY; PRT; 270 AA.
 DT 01-NOV-1998 (TRENBLrel. 08, Created).
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE RAS ASSOCIATION (RALGDS/AF-6) domain family 1 protein isoform 1C
 DE (RALGDS/AF-6) (Putative tumor suppressor protein).
 GN RAS321 OR RAS32.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Burbee D., Forgacs E., White M.A., Lerman M., Minna J.D.;
 RT "RAS321, a putative RAS effector and tumor suppressor from the human
 RT 3p21.3 critical region, contains alternatively spliced messages from
 RT two promoters.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Dammann R., Li C., Bates S., Pfeifer G.P.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Dammann R., Li C., Yoon J.-H., Chin P.L., Bates S., Pfeifer G.P.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF040703; AAC70910.2; -;
 DR EMBL: AF061836; AAC16001.1; -;
 DR EMBL: AF132676; AAC4475.1; -;
 DR InterPro: IPR000159; RA_domain.
 DR Pfam: PF00788; RA; 1.
 DR SMART: SM00314; RA; 1.
 SQ SEQUENCE 270 AA; 31226 MW; 2E7006EA16A38D25 CRC64;

Query Match 73.9%; Score 17; DB 4; Length 270;
 Best Local Similarity 25.0%; Pred. No. 7.4e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 13 WSTTSSG 20

RESULT 6

Q9WUF5 PRELIMINARY; PRT; 270 AA.
 AC Q9WUF5;
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE 123F2 protein (RAS ASSOCIATION domain family 1 isoform C) (RAS
 DE ASSOCIATION (RALGDS/AF-6) domain family 1).
 GN RAS321.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;

RA Duh F.-M., Minna J.D., Lerman M.I.;
 RT "Mouse ortholog of the human 123F2 gene";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvImJ;
 RA Dammann R., Pfeifer G.P.;
 RT "A mouse locus containing the ortholog of the human RAS321 tumor
 RT suppressor gene";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF132851; AAD30061.1; -;
 DR EMBL: AF333027; AAK21201.1; -;
 DR EMBL: BC002173; AAH02173.1; -;
 DR MGD; MGI:1928386; RAS321.
 DR InterPro: IPR000159; RA_domain.
 DR Pfam: PF00788; RA; 1.
 DR SMART: SM00314; RA; 1.
 SQ SEQUENCE 270 AA; 31193 MW; 6E11302FA93EC0BC CRC64;

Query Match 73.9%; Score 17; DB 11; Length 270;
 Best Local Similarity 25.0%; Pred. No. 7.4e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 13 WSTTSSG 20

RESULT 7

O8GLP9 PRELIMINARY; PRT; 274 AA.
 AC O8GLP9;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Cher.
 GN Cher.
 OS Aeromonas hydrophila.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Aeromonadales;
 CC Aeromonadaceae; Aeromonas.
 OX NCBI_TaxID=644;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tomas J.M., Altarrriba M.;
 RT "A polar flagella operon (flg) of Aeromonas hydrophila contains genes
 RT required for lateral flagella expression";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY129558; AAN08636.1; -;
 SQ SEQUENCE 274 AA; 30645 MW; 611537BDE3FFF093 CRC64;

Query Match 73.9%; Score 17; DB 2; Length 274;
 Best Local Similarity 25.0%; Pred. No. 7.5e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

Db 106 WSAASSG 113

RESULT 8

O8GM67 PRELIMINARY; PRT; 275 AA.
 ID O8GM67;
 AC O8GM67;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Cher.
 GN Cher.
 OS Vibrio fischeri.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=668;
 RP [1]
 RC STRAIN=ES114;
 RT Deloney C.R., Wolfe A.J., Visick K.L.;
 RT "Role of chemotaxis in the Vibrio fischeri-Euprymna scolopes
 symbiosis";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY079167; AAL86626.1; -
 SQ SEQUENCE 275 AA; 30809 MW; 6459E0136487786B CRC64;

Query Match 73.9%; Score 17; DB 2; Length 275;
 Best Local Similarity 25.0%; Pred. No. 7.5e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 WXXXXXXG 8

DB 108 WSAASSG 115

RESULT 9

ID QPIM6 PRELIMINARY; PRT; 275 AA.
 AC QPIM6;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Response regulator for chemotaxis.
 GN CHER OR XAC2869.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=92829;
 RP [1]
 RC SEQUENCE FROM N.A.
 RL Nature 417:459-463(2002).
 DR STRAIN=506 / ATCC 13902 / XV 101;
 RX MEDLINE=2202145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 host specificities";
 RL Nature 417:459-463(2002).

DR EMBL: AE011928; AAM5714.1; -
 DR InterPro: IPR000780; ChEM_Metranf.
 DR InterPro: IPR001601; Methyltransf.
 DR Pfam: PF01739; ChEM; 1.
 DR Pfam: PF03705; ChEM; 1.
 DR PRINTS: PR00996; ChEMTFRASE.
 DR PROSITE: PS50123; ChEM; 1.
 KW Complete proteome.
 SQ SEQUENCE 275 AA; 31290 MW; 884C688AD01832F0 CRC64;

Query Match 73.9%; Score 17; DB 16; Length 275;
 Best Local Similarity 25.0%; Pred. No. 7.5e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 WXXXXXXG 8

DB 114 WSAASSTG 121

DB 114 WSAASSG 121

RESULT 10

ID Q8P7A7 PRELIMINARY; PRT; 275 AA.
 AC Q8P7A7;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Response regulator for chemotaxis.
 GN CHER OR XCC2704.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RP [1]
 RC SEQUENCE FROM N.A.
 RL Nature 417:459-463(2002).
 DR STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=2202145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A., L.P.,
 RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 host specificities";
 RL Nature 417:459-463(2002).
 DR EMBL: AE012383; AAM41976.1; -
 DR InterPro: IPR000780; ChEM_Metranf.
 DR InterPro: IPR001601; Methyltransf.
 DR Pfam: PF01739; ChEM; 1.
 DR Pfam: PF03705; ChEM; 1.
 DR PRINTS: PR00996; ChEMTFRASE.
 DR PROSITE: PS50123; ChEM; 1.
 KW Complete proteome.
 SQ SEQUENCE 275 AA; 31155 MW; 1FDCDD20336AB789 CRC64;

Query Match 73.9%; Score 17; DB 16; Length 275;
 Best Local Similarity 25.0%; Pred. No. 7.5e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 WXXXXXXG 8

DB 114 WSAASSTG 121

RESULT 11

ID Q8DFI4 PRELIMINARY; PRT; 275 AA.
 AC Q8DFI4;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Methylase of chemotaxis methyl-accepting protein ChEM.
 GN VW10227.
 OS Vibrio vulnificus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=672;
 RP [1]
 RC SEQUENCE FROM N.A.

Query Match 73.9%; Score 17; DB 16; Length 275;
 Best Local Similarity 25.0%; Pred. No. 7.5e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 WXXXXXXG 8

DB 114 WSAASSTG 121

RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RA Choy H.E.;
 RT "Complete genome sequence of *Vibrio vulnificus* CMCP6.";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE016797; AAC08764.1; -;
 KW Complete proteome.
 SQ SEQUENCE 275 AA; 30888 MW; B87667B6FB3F191B CRC64;

Query Match 73.9%; Score 17; DB 16; Length 275;
 Best Local Similarity 25.0%; Pred. No. 7.5e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8

Db 108 WSAASSG 115

RESULT 12

Q9AAJ9 ID Q9AAJ9 PRELIMINARY; PRT; 276 AA.
 AC Q9AAJ9;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE Chemotaxis protein methyltransferase CheR.
 GN CC0598.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=153892;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smith R.J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of *Caulobacter crescentus*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AE003734; AAK22584.1; -;
 DR HSSP: P07801; 1AF7.
 DR TIGR: CC0598; -;
 DR InterPro: IPR000780; CheR_Metranf.
 DR InterPro: IPR001601; Methyltransf.
 DR Pfam: PF01739; CheR; 1.
 DR PRINTS: PR00996; CHERMTFRASE.
 DR SMART: SM00138; MetTc; 1.
 DR PROSITE: PS50123; CHER; 1.
 KW Transferase; Methyltransferase; Complete proteome.
 SQ SEQUENCE 276 AA; 30570 MW; B99D30457C9E2125 CRC64;

Query Match 73.9%; Score 17; DB 16; Length 276;
 Best Local Similarity 25.0%; Pred. No. 7.5e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8

Db 103 WSAASTG 110

RESULT 13

Q8F114 ID Q8F114 PRELIMINARY; PRT; 283 AA.
 AC Q8F114;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Hypothetical protein.

GN LA3325.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE011492; RAN50523.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 283 AA; 31176 MW; 1D303BD0BFFA129 CRC64;

Query Match 73.9%; Score 17; DB 16; Length 283;
 Best Local Similarity 25.0%; Pred. No. 7.7e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8

Db 251 WSSSTASG 258

RESULT 14

Q8XCF8 ID Q8XCF8 PRELIMINARY; PRT; 286 AA.
 AC Q8XCF8;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Response regulator for chemotaxis, protein glutamate methyltransferase.
 DE methyitransferase.
 GN CHER OR Z2938 OR ECS2594.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Rose D.J., Mayhew G.F., Hackett J., Link S., Boutin A., Shao Y., Miller L.,
 RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 DR EMBL: AE005410; AAG56874.1; -;
 DR EMBL: AP002559; BAB36017.1; -;
 DR InterPro: IPR000780; CheR_Metranf.
 DR InterPro: IPR001601; Methyltransf.
 DR InterPro: IPR000051; SAM_bind.
 DR Pfam: PF01739; CheR; 1.
 DR Pfam: PF03705; CheR_N; 1.
 DR PRINTS: PR00996; CHERMTFRASE.
 DR SMART: SM00138; MetTc; 1.
 DR PROSITE: PS50123; CHER; 1.
 KW Complete proteome.
 SQ SEQUENCE 286 AA; 32839 MW; 9507A07DD7CAYE36 CRC64;

Query Match 73.9%; Score 17; DB 16; Length 286;

Best Local Similarity 25.0%; Pred. No. 7.8e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 120 WSAASTG 127

RESULT 15

Q8PIK6 PRELIMINARY; PRT; 286 AA.
AC Q8PIK6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein XAC2891.
GN XAC2891.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
CX NCBI_TaxID=92829;
FN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Montelero-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.F., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madsira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Seta J.A.D., Silva C., de Souza R.F.,
RA Spinoia L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
DR EMBL; AB011931; RAMS7736.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 286 AA; 31122 MW; 5A42AE9D1D381578 CRC64;

Query Match 73.3%; Score 17; DB 16; Length 286;
Best Local Similarity 25.0%; Pred. No. 7.8e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 60 WTSSAAG 67

Search completed: August 16, 2003, 14:33:30
Job time : 100 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:31:20 ; Search time 83 Seconds
(without alignments)
15.299 Million cell updates/sec

Title: US-09-185-908-1

Perfect score: 23

Sequence: 1 WXXXXXX 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	17	73.9	63	AAU43896	Propionibacterium
2	17	73.9	93	AAU21120	Human novel foetal
3	17	73.9	100	AAU44311	Propionibacterium
4	17	73.9	109	AAU30877	Novel human secret
5	17	73.9	129	AAU62174	Propionibacterium
6	17	73.9	145	AAU54416	Zea mays protein f
7	17	73.9	150	AAU54613	Zea mays protein f
8	17	73.9	174	ABR89383	Human polypeptide
9	17	73.9	209	ABG29051	Novel human diagno

10	17	73.9	256	23	ABB91973	Herbicidally activ
11	17	73.9	263	21	AAG26200	Arabidopsis thalia
12	17	73.9	263	21	AAG38398	Arabidopsis thalia
13	17	73.9	270	23	ABG71309	Human tumour suppr
14	17	73.9	270	23	AAU98088	Human RassF1 prev
15	17	73.9	270	23	AAU98470	Human ras effector
16	17	73.9	274	21	AAG38397	Arabidopsis thalia
17	17	73.9	275	21	AAG26199	Arabidopsis thalia
18	17	73.9	299	22	AAG90486	C glutamicum prote
19	17	73.9	354	24	ABJ26660	Human protein modi
20	17	73.9	432	24	ABU11362	Protein encoded by
21	17	73.9	495	22	ABG07440	Novel human diagno
22	17	73.9	527	19	AAW75776	Mycobacterium tube
23	17	73.9	527	19	AAW48359	Mycobacterium bovi
24	17	73.9	562	22	ABG29116	Novel human diagno
25	17	73.9	593	17	AAU92278	Human kidney PTH/P
26	17	73.9	593	20	AAW73317	Human Parathyroid
27	17	73.9	593	22	ABG56385	Non-endogenous hum
28	17	73.9	593	22	AAU71876	Human PTHR seven t
29	17	73.9	593	24	ABP81872	Human parathyroid
30	17	73.9	595	24	ABG73825	Canine parathyroid
31	17	73.9	614	13	AAU27707	Human kidney PTH/P
32	17	73.9	1130	22	ABG02249	Novel human diagno
33	16	69.6	17	23	ABU07407	xyNA fragment in p
34	16	69.6	20	21	AAU78408	Human papillomavir
35	16	69.6	20	23	ABU07406	xyNA fragment in p
36	16	69.6	32	21	ABU38243	Human secreted pro
37	16	69.6	33	19	AAW44053	Mutant hCS peptide
38	16	69.6	51	21	AAU76131	Human secreted pro
39	16	69.6	51	22	ABG48652	Human liver peptid
40	16	69.6	51	22	ABG28633	Peptide #1284 enco
41	16	69.6	51	22	ABG3817	Peptide #1323 enco
42	16	69.6	51	22	ABU19259	Protein #1258 enco
43	16	69.6	51	22	AAU54583	Human brain expres
44	16	69.6	51	22	AAU66989	Human bone marrow
45	16	69.6	51	22	AAU14849	Peptide #1283 enco

ALIGNMENTS

RESULT 1

AAU43896
ID AAU43896 standard; Protein; 63 AA.

AC AAU43896;

DT 13-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #4792.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

OS WO200181581-A2.

PN 01-NOV-2001.

PD 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR WPI: 2001-616774/71.
 DR N-PSDB; AA859521.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -
 XX
 XX
 PS Example 1; SEQ ID NO 5091; 1069pp; English.
 XX
 CC Sequences AAU99105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 63 AA;
 SQ
 Query Match 73.9%; Score 17; DB 22; Length 63;
 Best Local Similarity 25.0%; Pred. No. 5.2e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 WXXXXXXG 8
 Db 56 WAASSTSG 63
 RESULT 2
 AAU21120
 ID AAU21120 standard; Protein; 93 AA.
 XX
 AC AAU21120;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human novel foetal antigen, SEQ ID NO 1364.
 XX
 KW Human; foetal tissue antigen; antiinflammatory; neuroprotective;
 KW immunomodulator; cardiovascular; cytostatic; nephrotoxic;
 KW cardiovascular; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; breast neoplasm; cancer;
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
 KW cerebral ischaemia; angiogenesis; nervous system disorder;
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;
 KW wound healing; epithelial cell proliferation; food additive.
 XX
 OS Homo sapiens.
 XX
 XX WO200155312-A2.
 PN
 XX
 XX 02-AUG-2001.
 PD
 XX
 PF 17-JAN-2001; 2001WO-US01321.
 XX
 XX 31-JAN-2000; 2000US-0173065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR

PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.

PR	13-OCT-2000;	2000US-0235937.	CC	The invention relates to novel nucleic acids encoding novel human foetal
PR	20-OCT-2000;	2000US-0240960.	CC	antigens. The nucleic acids and proteins are used to prevent, treat (e.g.
PR	20-OCT-2000;	2000US-0241221.	CC	by gene therapy) or ameliorate a medical condition in e.g. humans, mice,
PR	20-OCT-2000;	2000US-0241785.	CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They
PR	20-OCT-2000;	2000US-0241786.	CC	are also used in diagnosing a pathological condition or susceptibility
PR	20-OCT-2000;	2000US-0241787.	CC	to a pathological condition. The antibodies to the antigens can also
PR	20-OCT-2000;	2000US-0241808.	CC	be used in alleviating symptoms associated with the disorders and in
PR	20-OCT-2000;	2000US-0241809.	CC	diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
PR	20-OCT-2000;	2000US-0241826.	CC	immunosorbent assays (ELISA). Disorders which are diagnosed or treated
PR	01-NOV-2000;	2000US-0244617.	CC	include autoimmune diseases e.g. rheumatoid arthritis,
PR	08-NOV-2000;	2000US-0246474.	CC	hyperproliferative disorders e.g. neoplasms of the breast or liver,
PR	08-NOV-2000;	2000US-0246475.	CC	cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
PR	08-NOV-2000;	2000US-0246476.	CC	e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
PR	08-NOV-2000;	2000US-0246477.	CC	Alzheimer's disease, infections caused by bacteria, viruses and fungi
PR	08-NOV-2000;	2000US-0246478.	CC	and ocular disorders e.g. corneal infection. The polypeptides can also
PR	08-NOV-2000;	2000US-0246523.	CC	be used to aid wound healing and epithelial cell proliferation, to
PR	08-NOV-2000;	2000US-0246524.	CC	prevent skin aging due to sunburn, to maintain organs before
PR	08-NOV-2000;	2000US-0246525.	CC	transplantation, for supporting cell culture of primary tissues, to
PR	08-NOV-2000;	2000US-0246526.	CC	regenerate tissues and in chemotaxis. The polypeptides can also be used
PR	08-NOV-2000;	2000US-0246527.	CC	as a food additive or preservative to increase or decrease storage
PR	08-NOV-2000;	2000US-0246528.	CC	capabilities, fat content, lipid, protein, carbohydrate, vitamins,
PR	08-NOV-2000;	2000US-0246532.	CC	minerals, cofactors and other nutritional components. Numerous
PR	08-NOV-2000;	2000US-0246609.	CC	examples of diseases and disorders treated by the nucleic acids and
PR	08-NOV-2000;	2000US-0246610.	CC	proteins are given in the specification. The present sequence
PR	08-NOV-2000;	2000US-0246611.	CC	
PR	08-NOV-2000;	2000US-0246613.	CC	
PR	17-NOV-2000;	2000US-0249207.	CC	
PR	17-NOV-2000;	2000US-0249208.	CC	
PR	17-NOV-2000;	2000US-0249209.	CC	
PR	17-NOV-2000;	2000US-0249210.	CC	
PR	17-NOV-2000;	2000US-0249211.	CC	
PR	17-NOV-2000;	2000US-0249212.	CC	
PR	17-NOV-2000;	2000US-0249213.	CC	
PR	17-NOV-2000;	2000US-0249214.	CC	
PR	17-NOV-2000;	2000US-0249215.	CC	
PR	17-NOV-2000;	2000US-0249216.	CC	
PR	17-NOV-2000;	2000US-0249217.	CC	
PR	17-NOV-2000;	2000US-0249218.	CC	
PR	17-NOV-2000;	2000US-0249244.	CC	
PR	17-NOV-2000;	2000US-0249245.	CC	
PR	17-NOV-2000;	2000US-0249264.	CC	
PR	17-NOV-2000;	2000US-0249265.	CC	
PR	17-NOV-2000;	2000US-0249267.	CC	
PR	17-NOV-2000;	2000US-0249297.	CC	
PR	17-NOV-2000;	2000US-0249299.	CC	
PR	17-NOV-2000;	2000US-0249300.	CC	
PR	01-DEC-2000;	2000US-0250160.	CC	
PR	01-DEC-2000;	2000US-0250391.	CC	
PR	05-DEC-2000;	2000US-0251030.	CC	
PR	05-DEC-2000;	2000US-0251988.	CC	
PR	05-DEC-2000;	2000US-0256719.	CC	
PR	06-DEC-2000;	2000US-0251479.	CC	
PR	08-DEC-2000;	2000US-0251856.	CC	
PR	08-DEC-2000;	2000US-0251868.	CC	
PR	08-DEC-2000;	2000US-0251869.	CC	
PR	08-DEC-2000;	2000US-0251989.	CC	
PR	08-DEC-2000;	2000US-0251990.	CC	
PR	11-DEC-2000;	2000US-0254097.	CC	
PR	05-JAN-2001;	2001US-0259678.	CC	
XX			CC	
PA			CC	(HUMA-) HUMAN GENOME SCI INC.
XX			CC	
PI	Rosen CA, Barash SC, Ruben SM;		CC	
XX			CC	
DR	WPI; 2001-488782/53.		CC	
DR	N-PSDB; AAS33940.		CC	
XX			CC	
XX			CC	
PT	New polynucleotides and polypeptides for diagnosing, treating,		CC	
PT	preventing or progressing e.g. diseases or disorders of the nervous,		CC	
PT	musculoskeletal, excretory, gastrointestinal, reproductive, and		CC	
PT	respiratory systems -		CC	
XX				

Example 1; SEQ ID No 6106; 1069pp; English.

Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 100 AA;

Query Match 73.9%; Score 17; DB 22; Length 100;
Best Local Similarity 25.0%; Pred. No. 7.6e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 87 WTTSSSG 94

RESULT 4
AAU30877
ID AAU30877 standard; Protein; 109 AA.

XX AC AAU30877;

XX 18-DEC-2001 (first entry)

XX Novel human secreted protein #1368.

XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US08656.

XX 18-APR-2000; 2000US-0552929.

XX 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -

XX Claim 20; Page 365; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The

CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated

CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.

XX Sequence 109 AA;

Query Match 73.9%; Score 17; DB 22; Length 109;
Best Local Similarity 25.0%; Pred. No. 8.1e+03;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
DB 19 WASSASSG 26

RESULT 5
AAU62174
ID AAU62174 standard; Protein; 129 AA.

XX AC AAU62174;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #23070.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L' Maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59623.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

XX Example 1; SEQ ID No 23369; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 129 AA;

Query Match 73.98; Score 17; DB 22; Length 129;
 Best Local Similarity 25.08; Pred. No. 9.2e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 WXXXXXXG 8

Db 88 WWTSSG 95

RESULT 6

AAG54416
 ID AAG54416 standard; Protein; 145 AA.

AC AAG54416;

XX 18-OCT-2000 (first entry)

XX Zea mays protein fragment SEQ ID NO: 69382.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.

XX Zea mays subsp. mays.

XX EF1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 03-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127452.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.

```
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147202.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150586.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0159369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159322.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 28-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 73.9%; Score 17; DB 21; Length 145;
Best Local Similarity 25.0%; Pred. NO. let04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
Db 59 WTTTASAG 66

RESULT 7
AAG54613
ID AAG54613 standard; Protein; 150 AA.
XX
AC AAG54613;
XX
DF 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 69653.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
```

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136762.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149315.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 73.9%; Score 17; DB 21; Length 150;
 Best Local Similarity 25.0%; Pred. No. 1e+04;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 DB 64 WTTTASAG 71

RESULT 8
 ABB89383
 ID ABB89383 standard; Protein; 174 AA.
 XX
 AC ABB89383;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1759.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiact; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO2001190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US16450.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 XX WPI; 2002-122018/16.
 DR N-PSDB; ABL89792.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 PS Claim 11; SEQ ID NO 1759; 2081pp + Sequence Listing; English.
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 174 AA;

Query Match 73.9%; Score 17; DB 23; Length 174;
 Best Local Similarity 25.0%; Pred. No. 1.2e+04;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 DB 97 WSSAAAG 104

RESULT 9
 ABG29051
 ID ABG29051 standard; Protein; 209 AA.
 XX
 AC ABG29051;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #29042.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO2001175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS93238.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 XX diagnostics, forensics, gene mapping, identification of mutations
 XX responsible for genetic disorders or other traits and to assess
 XX biodiversity -
 XX
 PS Claim 20; SEQ ID NO 59410; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 209 AA;

Query Match 73.9%; Score 17; DB 22; Length 209;
 Best Local Similarity 25.0%; Pred. No. 1.4e+04;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

DB 14 WSAATSG 21

RESULT 10

ABB91973

ID ABB91973 standard; Protein; 256 AA.

XX AC ABB91973;

XX DT 31-MAY-2002 (first entry)

XX DE Herbicidally active polypeptide SEQ ID NO 1184.

XX KW Herbicidal; plant; agriculture; herbicide.

XX OS Arabidopsis thaliana.

XX PN WC200210210-A2.

XX PD 07-FEB-2002.

XX EF 28-AUG-2001; 2001WO-EP09892.

XX PR 28-AUG-2001; 2001WO-EP09892.

XX PA (FARB) BAYER AG.

XX PI Tietjen K, Weidner M;

XX DR WPI; 2002-269010/31.

XX PT Identifying plant target proteins for herbicidally active compounds,

PT comprising aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant

PT organisms -

XX PS Claim 5; SEQ ID NO 1184; 261pp + Sequence Listing; English.

XX CC The invention relates to identifying target proteins

CC (ABB90790-ABB94016) for herbicidally active compounds, comprising

CC aligning and comparing nucleic acid or amino acid sequences from plant

CC with nucleic acid or amino acid sequences from non-plant organisms using

CC suitable search parameters, where plant sequences having an B-value

CC greater by a factor of 3 than the B-value of most similar non-plant

CC sequences are selected. The polypeptides or nucleic acids encoding them

CC are useful for identifying modulators. The identified modulators are

XX useful as herbicides.

XX SQ Sequence 256 AA;

Query Match 73.9%; Score 17; DB 23; Length 256;

Best Local Similarity 25.0%; Pred. No. 1.6e+04;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

DB 222 WATATATG 229

RESULT 11

AAG26200

ID AAG26200 standard; Protein; 263 AA.

XX AC AAG26200;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 30568.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126284.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145132.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 23-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 73.9%; Score 17; DB 21; Length 263;
Best Local Similarity 25.0%; Pred. No. 1.6e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
Db 229 WATATATG 236

RESULT 12
AAG38398
ID AAG38398 standard; Protein; 263 AA.

XX AAG38398;
XX AC
XX DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 47365.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000. 99US-0142055.
PD 06-JUL-1999; 99US-0142390.
XX 25-FEB-2000; 2000EP-0301439. PR 08-JUL-1999; 99US-0142803.
XX 25-FEB-1999; 99US-0121825. PR 09-JUL-1999; 99US-0142920.
PD 05-MAR-1999; 99US-0123180. PR 12-JUL-1999; 99US-0142977.
XX 09-MAR-1999; 99US-0123546. PR 13-JUL-1999; 99US-0143542.
PD 23-MAR-1999; 99US-0125788. PR 14-JUL-1999; 99US-0143624.
PD 23-MAR-1999; 99US-0126864. PR 15-JUL-1999; 99US-0144005.
PD 29-MAR-1999; 99US-0126785. PR 16-JUL-1999; 99US-0144085.
PD 01-APR-1999; 99US-0127462. PR 18-JUL-1999; 99US-0144086.
PD 06-APR-1999; 99US-0128234. PR 19-JUL-1999; 99US-0144325.
PD 08-APR-1999; 99US-0128714. PR 19-JUL-1999; 99US-0144331.
PD 16-APR-1999; 99US-0129845. PR 19-JUL-1999; 99US-0144333.
PD 19-APR-1999; 99US-0130077. PR 19-JUL-1999; 99US-0144334.
PD 21-APR-1999; 99US-0130449. PR 19-JUL-1999; 99US-0144335.
PD 23-APR-1999; 99US-0130510. PR 20-JUL-1999; 99US-0144352.
PD 28-APR-1999; 99US-0130891. PR 20-JUL-1999; 99US-0144632.
PD 30-APR-1999; 99US-0131449. PR 20-JUL-1999; 99US-0144884.
PD 04-MAY-1999; 99US-0132048. PR 21-JUL-1999; 99US-0144814.
PD 05-MAY-1999; 99US-0132407. PR 21-JUL-1999; 99US-0145086.
PD 06-MAY-1999; 99US-0132485. PR 21-JUL-1999; 99US-0145088.
PD 06-MAY-1999; 99US-0132486. PR 22-JUL-1999; 99US-0145085.
PD 07-MAY-1999; 99US-0132487. PR 22-JUL-1999; 99US-0145087.
PD 11-MAY-1999; 99US-0132863. PR 22-JUL-1999; 99US-0145089.
PD 14-MAY-1999; 99US-0134256. PR 23-JUL-1999; 99US-0145145.
PD 14-MAY-1999; 99US-0134218. PR 23-JUL-1999; 99US-0145218.
PD 14-MAY-1999; 99US-0134219. PR 23-JUL-1999; 99US-0145224.
PD 14-MAY-1999; 99US-0134221. PR 26-JUL-1999; 99US-0145276.
PD 18-MAY-1999; 99US-0134370. PR 27-JUL-1999; 99US-0145913.
PD 19-MAY-1999; 99US-0134768. PR 27-JUL-1999; 99US-0145918.
PD 20-MAY-1999; 99US-0134941. PR 27-JUL-1999; 99US-0145919.
PD 21-MAY-1999; 99US-0135124. PR 28-JUL-1999; 99US-0145951.
PD 21-MAY-1999; 99US-0135353. PR 02-AUG-1999; 99US-0146386.
PD 25-MAY-1999; 99US-0136021. PR 02-AUG-1999; 99US-0146388.
PD 27-MAY-1999; 99US-0136392. PR 03-AUG-1999; 99US-0146389.
PD 28-MAY-1999; 99US-0136782. PR 04-AUG-1999; 99US-0147038.
PD 01-JUN-1999; 99US-0137222. PR 04-AUG-1999; 99US-0147204.
PD 03-JUN-1999; 99US-0137528. PR 04-AUG-1999; 99US-0147302.
PD 04-JUN-1999; 99US-0137724. PR 05-AUG-1999; 99US-0147192.
PD 07-JUN-1999; 99US-0137724. PR 06-AUG-1999; 99US-0147260.
PD 08-JUN-1999; 99US-0138094. PR 06-AUG-1999; 99US-0147303.
PD 10-JUN-1999; 99US-0138540. PR 09-AUG-1999; 99US-0147493.
PD 10-JUN-1999; 99US-0138847. PR 09-AUG-1999; 99US-0147935.
PD 14-JUN-1999; 99US-0139119. PR 10-AUG-1999; 99US-0147935.
PD 16-JUN-1999; 99US-0139452. PR 11-AUG-1999; 99US-0148171.
PD 17-JUN-1999; 99US-0139453. PR 12-AUG-1999; 99US-0148319.
PD 18-JUN-1999; 99US-0139452. PR 13-AUG-1999; 99US-0148341.
PD 18-JUN-1999; 99US-0139454. PR 13-AUG-1999; 99US-0148565.
PD 18-JUN-1999; 99US-0139455. PR 16-AUG-1999; 99US-0148684.
PD 18-JUN-1999; 99US-0139456. PR 17-AUG-1999; 99US-0149368.
PD 18-JUN-1999; 99US-0139457. PR 18-AUG-1999; 99US-0149175.
PD 18-JUN-1999; 99US-0139458. PR 18-AUG-1999; 99US-0149426.
PD 18-JUN-1999; 99US-0139459. PR 20-AUG-1999; 99US-0149722.
PD 18-JUN-1999; 99US-0139460. PR 20-AUG-1999; 99US-0149723.
PD 18-JUN-1999; 99US-0139461. PR 20-AUG-1999; 99US-0149929.
PD 18-JUN-1999; 99US-0139462. PR 23-AUG-1999; 99US-0149902.
PD 18-JUN-1999; 99US-0139463. PR 23-AUG-1999; 99US-0149930.
PD 18-JUN-1999; 99US-0139750. PR 25-AUG-1999; 99US-0150566.
PD 21-JUN-1999; 99US-0139763. PR 26-AUG-1999; 99US-0150884.
PD 22-JUN-1999; 99US-0139817. PR 27-AUG-1999; 99US-0151065.
PD 23-JUN-1999; 99US-0139899. PR 27-AUG-1999; 99US-0151066.
PD 23-JUN-1999; 99US-0140353. PR 30-AUG-1999; 99US-0151080.
PD 24-JUN-1999; 99US-0140354. PR 31-AUG-1999; 99US-0151303.
PD 28-JUN-1999; 99US-0140823. PR 01-SEP-1999; 99US-0151438.
PD 29-JUN-1999; 99US-0140991. PR 07-SEP-1999; 99US-0151930.
PD 30-JUN-1999; 99US-0141287. PR 10-SEP-1999; 99US-0152363.
PD 01-JUL-1999; 99US-0141842. PR 13-SEP-1999; 99US-0153070.
PD 01-JUL-1999; 99US-0141842. PR 15-SEP-1999; 99US-0153758.
PD 01-JUL-1999; 99US-0141842. PR 16-SEP-1999; 99US-0154018.
PD 01-JUL-1999; 99US-0141842. PR 20-SEP-1999; 99US-0154039.
PD 01-JUL-1999; 99US-0141842. PR 20-SEP-1999; 99US-0154779.

PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159594.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 73.9%; Score 17; DB 21; Length 263;
 Best Local Similarity 25.0%; Pred. No. 1.6e+04;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 DB 229 WATATATG 236

RESULT 13
 ABG71309
 ID ABG71309 standard; Protein; 270 AA.

XX AC ABG71309;

XX DT 19-DEC-2002 (first entry)

XX DE Human tumour suppressor RASSF1.C.

XX KW Human; tumour suppressor; RASSF1.C; cancer; breast cancer;
 XX KW DNA methylation; lung cancer; kidney cancer; ovarian cancer;
 XX KW head and neck cancer; melanoma; chromosome 3p21.3.

XX OS Homo sapiens.

XX PN US2002098530-A1.

XX PD 25-JUL-2002.

XX PF 30-MAR-2001; 2001US-0821803.

XX PR 30-MAR-2000; 2000US-193268P.

XX PN

PA (CITY) CITY OF HOPE.

XX PI Pfeifer GP, Dammann R;

XX WP1; 2002-690479/74.

DR N-PSDB; ABS55576.

XX Novel tumor suppressor gene, termed RASSF1, useful for the diagnosis of
 predisposition to cancer by analyzing its methylation status,
 heterozygosity or mutation

Claim 4; Fig 1C; 57pp; English.

The invention relates to an isolated tumour suppressor gene coding for
 splice variant RASSF1.A, RASSF1.B or RASSF1.C protein or its complement,
 or a DNA molecule which hybridises under stringent conditions to them.
 Also included are naturally occurring mutants of RASSF1.A, detecting (M1)
 a methylated RASSF1 gene, non-expressed RASSF1 gene or an alteration in
 RASSF1 where the methylation, non-expression or alteration is associated
 with cancer in a human, by analysing an RASSF1 gene or an RASSF gene
 expression product from a tissue or body fluid of the human;
 administering RASSF1 agonists to treat cancer, a RASSF1 non-human
 transgenic animal, a cell line from the transgenic animal, and
 screening for cancer therapeutics/drug candidates useful in treating
 cancer resulting from a methylated or a mutation in RASSF1.
 (M1) is useful for detecting methylated RASSF1 gene, which is
 useful for determining whether a human subject has or is at risk for
 developing cancer. The method involves detecting the methylation or
 non-expression of the gene or the presence or absence of a genetic
 polymorphism as in the RASSF1 gene of the subject, where the
 methylation or non-expression or the presence of the genetic
 polymorphism identifies a subject that has or is at risk for developing
 cancer. The mutants are useful for screening for drug candidates useful
 in treating cancer resulting from the RASSF1 gene. Analysis of the RASSF1
 gene is useful in the diagnosis of predisposition to cancer, including
 lung, breast, kidney, ovarian, head and neck cancer and melanoma. The
 association between the RASSF1 gene and cancer permits the early
 presymptomatic screening of individuals to identify those at risk for
 developing cancer. RASSF1 protein is useful for identifying agonists of
 the biological function of an RASSF1 protein. RASSF1, its encoding
 nucleic acids, antibodies and compounds identified by the screening
 assays are useful for treating cancer. The gene for RASSF1 is located
 on chromosome 3p21p3. The present sequence represents the RASSF1 splice
 variant protein RASSF1.C.

XX SQ Sequence 270 AA;

Query Match 73.9%; Score 17; DB 23; Length 270;

Best Local Similarity 25.0%; Pred. No. 1.7e+04;

Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

DB 13 WSRRTSSG 20

RESULT 14

AAU98088

ID AAU98088 standard; Protein; 270 AA.

XX AC AAU98088;

XX DT 24-SEP-2002 (first entry)

XX DE Human RASSF1 prey protein sequence.

XX KW Human; beta TrCP; bTrCP; protein-protein interaction complex; Ras SFL;
 KW drug screening; selected interacting domain; SID; tumour; gene therapy;
 KW prey protein; cytostatic.

XX OS Homo sapiens.

XX PN WO200250261-A2.

```

XX PD 27-JUN-2002.
XX PF
XX PS
XX PS Claim 3; Fig 2; 96pp; English.
XX CC
XX CC The invention describes an isolated nucleic acid (I) encoding the
XX CC Ras effector with tumour suppressor activity and the ability to induce
XX CC apoptosis, Minnl. A vector expressing (I) is useful for treating a human
XX CC subject with a solid tumour (e.g., ovarian tumour) where at least one
XX CC mutation in a Ras-family gene results in increased Ras signalling
XX CC activity and reduced levels of a Minnl product relative to non-tumour
XX CC tissue of like origin. Delivery of the vector is through liposome-DNA
XX CC complexes and recombinant viruses or by systemic delivery and local
XX CC delivery particularly surgical delivery, implantation or injection. The
XX CC recombinant virus comprises operably linked recombinant nucleotide
XX CC sequences comprising a suitable promoter sequence and viral sequences,
XX CC where the viral sequences are adenovirus, adeno-associated virus,
XX CC retrovirus, herpes virus, vaccinia virus and Moloney virus sequences. An
XX CC antibody (II) against the protein encoded by (I) is useful for detecting
XX CC a Minnl polypeptide in a sample and detecting a Minnl polypeptide in an
XX CC array of tissue samples, where the tissue array comprises more than 100
XX CC tissue samples from normal and tumour tissues. The method also comprises
XX CC determining the cell type in the tissue sample that exhibits the
XX CC antigen-antibody complex. This is the amino acid sequence of the human
XX CC ras effector and tumour suppressor Minnl.
XX SQ Sequence 270 AA;
XX
XX Query Match 73.98; Score 17; DB 23; Length 270;
XX Best Local Similarity 25.0%; Pred. NO. 1.7e+04;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 1 WXXXXXXG 8
XX |
XX Db 13 WSSWSSG 20
XX
XX Search completed: August 16, 2003, 14:34:58
XX Job time : 85 secs

```

```

XX PD 27-JUN-2002.
XX PF
XX PS
XX PS Claim 1; Page 47; 84pp; English.
XX CC
XX CC The present invention relates to a new complex of protein-protein
XX CC interaction between beta-CP (not defined in specification) and Ras Sfl.
XX CC the protein-protein complex of the invention is useful for screening
XX CC drugs or agents that modulate interaction of proteins. In particular,
XX CC the protein complex is useful for identifying the Selected Interacting
XX CC Domains (SID). The modulating compounds detected can be used for
XX CC treating tumours. The polynucleotides encoding the protein complex may
XX CC be used in gene therapy. The present amino acid sequence represents the
XX CC human RasSfl protein that was used in the methods of the invention as
XX CC a prey protein.
XX SQ Sequence 270 AA;
XX
XX Query Match 73.98; Score 17; DB 23; Length 270;
XX Best Local Similarity 25.0%; Pred. NO. 1.7e+04;
XX Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX QY 1 WXXXXXXG 8
XX |
XX Db 13 WSSWSSG 20
XX
XX RESULT 15
XX AAU98470
XX ID AAU98470 standard; Protein; 270 AA.
XX AC AAU98470;
XX XX
XX XX 24-SEP-2002 (first entry)
XX XX Human ras effector and tumour suppressor Minnl.
XX DE
XX DE Ras effector; tumour suppressor; solid tumour; ovarian tumour;
XX KW apoptosis; Ras-family gene; Ras signalling activity; Minnl.
XX XX
XX OS Homo sapiens.
XX XX WO200246223-A2.
XX PN
XX PD 13-JUN-2002.
XX XX
XX PF 07-DEC-2001; 2001WO-US48514.
XX XX
XX PR 07-DEC-2000; 2000US-251971P.
XX XX
XX XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PA Clark G, Vos M;
XX XX
XX XX WPI; 2002-527910/56.
XX DR N-PSDB; ABR65817.
XX XX

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:33:39 ; Search time 55 Seconds
(without alignments)
19,035 Million cell updates/sec

Title: US-09-185-908-1

Perfect score: 23

Sequence: 1 WXXXXXXG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 492763 seqs, 131003257 residues

Total number of hits satisfying chosen parameters: 492763

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	73.9	270	9	US-09-821-803A-6
2	17	73.9	270	15	US-10-023-530-4
3	17	73.9	290	15	US-10-156-761-12952
4	17	73.9	299	10	US-09-738-626-4240
5	17	73.9	527	15	US-10-176-640-3
6	17	73.9	527	15	US-10-178-495-3
7	17	73.9	593	10	US-09-943-446-9
8	17	73.9	593	15	US-10-225-567A-229
9	17	73.9	595	10	US-09-943-446-6
10	17	73.9	964	12	US-10-017-161-710
11	16	69.6	50	11	US-09-948-783-125
12	16	69.6	51	9	US-09-864-761-34557
13	16	69.6	51	11	US-09-892-877-124
14	16	69.6	54	9	US-09-864-761-42230
15	16	69.6	87	9	US-09-867-550-204

16	69.6	105	9	US-09-766-396-26	Sequence 26, Appl
17	69.6	105	12	US-10-335-125-2	Sequence 2, Appl
18	69.6	105	12	US-10-335-125-13	Sequence 13, Appl
19	69.6	105	14	US-10-062-375-26	Sequence 26, Appl
20	69.6	122	12	US-10-238-075-1559	Sequence 1559, Ap
21	69.6	134	9	US-09-764-853-760	Sequence 760, App
22	69.6	134	15	US-10-091-438-227	Sequence 227, App
23	69.6	135	12	US-10-137-870-380	Sequence 380, App
24	69.6	135	12	US-10-140-018-380	Sequence 380, App
25	69.6	135	12	US-10-140-021-380	Sequence 380, App
26	69.6	135	12	US-10-140-274-380	Sequence 380, App
27	69.6	135	12	US-10-140-471-380	Sequence 380, App
28	69.6	135	12	US-10-140-807-380	Sequence 380, App
29	69.6	135	12	US-10-140-922-380	Sequence 380, App
30	69.6	135	12	US-10-140-924-380	Sequence 380, App
31	69.6	135	12	US-10-140-926-380	Sequence 380, App
32	69.6	135	12	US-10-141-698-380	Sequence 380, App
33	69.6	135	12	US-10-141-702-380	Sequence 380, App
34	69.6	135	12	US-10-141-704-380	Sequence 380, App
35	69.6	135	12	US-10-142-421-380	Sequence 380, App
36	69.6	135	12	US-10-142-432-380	Sequence 380, App
37	69.6	135	12	US-10-142-767-380	Sequence 380, App
38	69.6	135	12	US-10-143-033-380	Sequence 380, App
39	69.6	135	12	US-10-144-994-380	Sequence 380, App
40	69.6	135	12	US-10-145-628-380	Sequence 380, App
41	69.6	135	12	US-10-145-631-380	Sequence 380, App
42	69.6	135	12	US-10-145-633-380	Sequence 380, App
43	69.6	135	12	US-10-145-746-380	Sequence 380, App
44	69.6	135	12	US-10-145-748-380	Sequence 380, App
45	69.6	135	12	US-10-145-823-380	Sequence 380, App

ALIGNMENTS

RESULT 1
US-09-821-803A-6
; Sequence 6, Application US/09821803A
; Patent No. US20020098530A1
; GENERAL INFORMATION:
; APPLICANT: Pfeiffer, Gerd P.
; APPLICANT: Damman, Reinhard
; TITLE OF INVENTION: Lung Cancer Tumor Suppressor Gene
; FILE REFERENCE: 1954-335-11
; CURRENT APPLICATION NUMBER: US/09/821,803A
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,268
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 270
; TYPE: PRI
; ORGANISM: Homo sapiens
US-09-821-803A-6

Query Match 73.9%; Score 17; DB 9; Length 270;
Best Local Similarity 25.0%; Pred No. 1.1e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 WXXXXXXG 8
Db 13 WSSVTSSTG 20

RESULT 2
US-10-023-530-4
; Sequence 4, Application US/10023530
; Publication No. US20030007956A1
; GENERAL INFORMATION:
; APPLICANT: LEGRAIN, Pierre
; APPLICANT: BENAROUS, Richard
; APPLICANT: BLOT, Guillaume

; APPLICANT: LASSOT, Irina
; TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TROP
; FILE REFERENCE: B4717A
; CURRENT APPLICATION NUMBER: US/10/023,530
; CURRENT FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 60/256,276
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Rasf1
; LOCATION: (1)..(270)
; OTHER INFORMATION: tumor suppressor
US-10-023-530-4

Query Match 73.9%; Score 17; DB 15; Length 270;
Best Local Similarity 25.0%; Pred. No. 1.1e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
Db 13 WSTTSSG 20

RESULT 3
US-10-156-761-12952
; Sequence 12952, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIRA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12952
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12952

Query Match 73.9%; Score 17; DB 15; Length 290;
Best Local Similarity 25.0%; Pred. No. 1.1e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
Db 57 WAAASAAG 64

RESULT 4
US-09-738-626-4240
; Sequence 4240, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIRO
; APPLICANT: TATEISHI, NAOHO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4240
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4240

Query Match 73.9%; Score 17; DB 10; Length 299;
Best Local Similarity 25.0%; Pred. No. 1.2e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
Db 114 WAAASAAG 121

RESULT 5
US-10-176-640-3
; Sequence 3, Application US/10176640
; Publication No. US20030023056A1
; GENERAL INFORMATION:
; APPLICANT: Anand, Naveen N.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: GENES ENCODING MYCOBACTERIAL PROTEINS ASSOCIATED WITH
; FILE REFERENCE: 1038-1231 MIS
; CURRENT APPLICATION NUMBER: US/10/176,640
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 08/677,970
; PRIOR FILING DATE: 1996-07-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-176-640-3

Query Match 73.9%; Score 17; DB 15; Length 527;
Best Local Similarity 25.0%; Pred. No. 1.6e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
Db 17 WSAATATG 24

RESULT 6
US-10-178-495-3
; Sequence 3, Application US/10178495
; Publication No. US20030088082A1
; GENERAL INFORMATION:
; APPLICANT: Anand, Naveen N.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: GENES ENCODING MYCOBACTERIAL PROTEINS ASSOCIATED WITH
; FILE REFERENCE: 1038-1231 MIS
; CURRENT APPLICATION NUMBER: US/10/178,495
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 08/677,970
; PRIOR FILING DATE: 1996-07-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-178-495-3

; FILE REFERENCE: 1038-1230 MIS
; CURRENT APPLICATION NUMBER: US/10/178,495
; CURRENT FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 08/677,970
; PRIOR FILING DATE: 1996-07-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-178-495-3

Query Match 73.9% Score 17; DB 15; Length 527;
Best Local Similarity 25.0%; Pred. No. 1.6e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 17 WSAATATG 24

RESULT 7
US-09-943-446-9
; Sequence 9, Application US/09943446
; Patent No. US2002014677A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Castleberry, Tessa A.
; APPLICANT: Lu, Bihong
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven L.
; TITLE OF INVENTION: Canine Parathyroid Hormone 1 Receptor
; CURRENT APPLICATION NUMBER: US/09/943,446
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: US 60/229,170
; PRIOR FILING DATE: 2000-08-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-943-446-9

Query Match 73.9% Score 17; DB 10; Length 593;
Best Local Similarity 25.0%; Pred. No. 1.8e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 69 WTSASTSG 76

RESULT 8
US-10-225-567A-229
; Sequence 229, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 229

; LENGTH: 593
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-229

Query Match 73.9% Score 17; DB 15; Length 593;
Best Local Similarity 25.0%; Pred. No. 1.8e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 69 WTSASTSG 76

RESULT 9
US-09-943-446-6
; Sequence 6, Application US/09943446
; Patent No. US2002014677A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Castleberry, Tessa A.
; APPLICANT: Lu, Bihong
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven L.
; TITLE OF INVENTION: Canine Parathyroid Hormone 1 Receptor
; FILE REFERENCE: PC10891AGPR
; CURRENT APPLICATION NUMBER: US/09/943,446
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: US 60/229,170
; PRIOR FILING DATE: 2000-08-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Canis Familiaris
US-09-943-446-6

Query Match 73.9% Score 17; DB 10; Length 595;
Best Local Similarity 25.0%; Pred. No. 1.8e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
|
DB 69 WASASTSG 76

RESULT 10
US-10-017-161-710
; Sequence 710, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 710
; LENGTH: 964
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-710

Query Match 73.9% Score 17; DB 12; Length 964;
Best Local Similarity 25.0%; Pred. No. 2.4e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

DB 241 WTSASTSG 248

RESULT 11

US-09-948-783-125
 ; Sequence 125, Application US/09948783
 ; Publication No. US20030100051A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et. al.
 ; TITLE OF INVENTION: 97 Human secreted proteins
 ; FILE REFERENCE: P2028P2
 ; CURRENT APPLICATION NUMBER: US/09/948,783
 ; CURRENT FILING DATE: 2001-09-10
 ; PRIOR APPLICATION NUMBER: 60/231,846
 ; PRIOR FILING DATE: 2000-09-11
 ; PRIOR APPLICATION NUMBER: 09/892,877
 ; PRIOR FILING DATE: 2001-06-28
 ; PRIOR APPLICATION NUMBER: 03/437,658
 ; PRIOR FILING DATE: 1999-11-10
 ; PRIOR APPLICATION NUMBER: PCT/US99/09847
 ; PRIOR FILING DATE: 1999-05-06
 ; PRIOR APPLICATION NUMBER: 60/085,093
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: 60/085,094
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: 60/085,105
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: 60/085,180
 ; PRIOR FILING DATE: 1998-05-12
 ; PRIOR APPLICATION NUMBER: 60/085,927
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,906
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,924
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,922
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,921
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,923
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,925
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,928
 ; PRIOR FILING DATE: 1998-05-18
 ; PRIOR APPLICATION NUMBER: 60/085,920
 ; PRIOR FILING DATE: 1998-05-18
 ; NUMBER OF SEQ ID NOS: 465
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 125
 ; LENGTH: 50
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-948-783-125

Query Match 69.6%; Score 16; DB 11; Length 50;
 Best Local Similarity 25.08; Pred. No. 6.3e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

DB 34 WTSMATG 41

RESULT 12

US-09-864-761-34557
 ; Sequence 34557, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
 ; FILE REFERENCE: Aemica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 4917
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 34557
 ; LENGTH: 51
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC004859.2
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.2
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.1
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.5
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.9
 ; US-09-864-761-34557

Query Match 69.6%; Score 16; DB 9; Length 51;
 Best Local Similarity 25.08; Pred. No. 6.4e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8

DB 16 WSSASGSG 23

RESULT 13

US-09-892-877-124
 ; Sequence 124, Application US/09892877
 ; Publication No. US20030077809A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et. al.
 ; TITLE OF INVENTION: 97 Human secreted proteins
 ; FILE REFERENCE: P2028P1
 ; CURRENT APPLICATION NUMBER: US/09/892,877
 ; CURRENT FILING DATE: 2001-06-28
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/437,658
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-10
 ; NUMBER OF SEQ ID NOS: 461
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 124
 ; LENGTH: 51
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (51)
 ; OTHER INFORMATION: Xaa equals stop translation
 US-09-892-877-124

Query Match 69.6%; Score 16; DB 11; Length 51;
 Best Local Similarity 25.0%; Pred. No. 6.4e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
 |
 Db 34 WTSSMATG 41

RESULT 14

US-09-864-761-42230
 ; Sequence 42230, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Aecomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263,6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 42230
 ; LENGTH: 54
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL117692.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.6
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 8.2
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 13
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 28
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 12
 ; OTHER INFORMATION: EST_HUMAN HIT: AA130933.1, EVALUATE 1.00e-04
 US-09-864-761-42230

Query Match 69.6%; Score 16; DB 9; Length 54;
 Best Local Similarity 25.0%; Pred. No. 6.6e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
 |
 Db 29 WSGTATSG 36

RESULT 15

US-09-867-550-204
 ; Sequence 204, Application US/09867550
 ; Patent No. US20020082206A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Leach, Martin D.
 ; APPLICANT: Mehraban, Fuad,
 ; APPLICANT: Conley, Pamela
 ; APPLICANT: Law, Debbie
 ; APPLICANT: Topper, James
 ; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a
 ; TITLE OF INVENTION: Thereby
 ; FILE REFERENCE: 21402-013 (Cura-313)
 ; CURRENT APPLICATION NUMBER: US/09/867,550
 ; CURRENT FILING DATE: 2001-09-20
 ; PRIOR APPLICATION NUMBER: USSN 60/208,427
 ; PRIOR FILING DATE: 2000-05-30
 ; NUMBER OF SEQ ID NOS: 2125
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 204
 ; LENGTH: 87
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-867-550-204

Query Match 69.6%; Score 16; DB 9; Length 87;
 Best Local Similarity 25.0%; Pred. No. 8.8e+03;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
 |
 Db 21 WSHSTSTG 28

Search completed: August 16, 2003, 14:43:28

Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2003, 14:31:50 ; Search time 377 Seconds
(without alignments)
18.469 Million cell updates/sec

Title: US-09-185-908-1
Perfect score: 23
Sequence: 1 WXXXXXG 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	17	73.9	31	20	US-09-617-682A-15802	Sequence 15802, A
2	17	73.9	63	1	PCT-US02-32727-5091	Sequence 5091, Ap

3	17	73.9	63	25	US-09-978-825-5091	Sequence 5091, Ap
4	17	73.9	63	26	US-10-057-498-5091	Sequence 5091, Ap
5	17	73.9	82	20	US-09-620-111B-3466	Sequence 3466, Ap
6	17	73.9	84	23	US-09-834-366-14776	Sequence 14776, A
7	17	73.9	84	31	US-60-197-873-14776	Sequence 14776, A
8	17	73.9	93	1	PCT-US01-01321-1364	Sequence 1364, Ap
9	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
10	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
11	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
12	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
13	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
14	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
15	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
16	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
17	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
18	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
19	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
20	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
21	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
22	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
23	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
24	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
25	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
26	17	73.9	98	29	US-10-316-253-118	Sequence 118, App
27	17	73.9	100	1	PCT-US02-32727-6106	Sequence 6106, Ap
28	17	73.9	100	25	US-09-978-825-6106	Sequence 6106, Ap
29	17	73.9	100	26	US-10-057-498-6106	Sequence 6106, Ap
30	17	73.9	104	30	US-10-437-963-124863	Sequence 124863,
31	17	73.9	105	30	US-10-437-963-178958	Sequence 178958,
32	17	73.9	106	30	US-10-437-963-119697	Sequence 119697,
33	17	73.9	109	1	PCT-US01-08656-7363	Sequence 7363, Ap
34	17	73.9	117	30	US-10-437-963-132266	Sequence 132266,
35	17	73.9	125	19	US-09-594-595B-977	Sequence 977, App
36	17	73.9	129	1	PCT-US02-32727-23369	Sequence 23369, A
37	17	73.9	129	25	US-09-978-825-23369	Sequence 23369, A
38	17	73.9	129	26	US-10-057-498-23369	Sequence 23369, A
39	17	73.9	129	30	US-10-437-963-142637	Sequence 142637,
40	17	73.9	130	21	US-09-708-427-19879	Sequence 19879, A
41	17	73.9	130	21	US-09-708-427-49929	Sequence 49929, A
42	17	73.9	130	21	US-09-708-427-55875	Sequence 55875, A
43	17	73.9	131	19	US-09-594-595B-976	Sequence 976, App
44	17	73.9	132	19	US-09-594-595B-975	Sequence 975, App
45	17	73.9	137	30	US-10-419-128-32654	Sequence 32654, A

ALIGNMENTS

RESULT 1

US-09-617-682A-15802
; Sequence 15802, Application US/09617682A
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypepti
; FILE REFERENCE: 2750-1063P
; CURRENT APPLICATION NUMBER: US/09/617,682A
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 16871
; SEQ ID NO 15802
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..31
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..31
; OTHER INFORMATION: Ceres Seq. ID 1424727
US-09-617-682A-15802

Query Match 73.9%; Score 17; DB 20; Length 31;
 Best Local Similarity 25.0%; Pred. No. 2.9e+04;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 |
 Db 16 WATTAAG 23

RESULT 2

PCT-US02-32727-5091
 ; Sequence 5091, Application PC/TUS0232727
 ; GENERAL INFORMATION:
 ; APPLICANT: Mitcham, Jennifer
 ; APPLICANT: Skeiky, Yasir
 ; APPLICANT: Persing, David
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Maisonneuve, Jean Francois
 ; APPLICANT: Zhang, Yanni
 ; APPLICANT: Wang, Siquing
 ; APPLICANT: Jen, Shyian
 ; APPLICANT: Lodes, Michael
 ; APPLICANT: Benson, Darin
 ; APPLICANT: Jones, Robert
 ; APPLICANT: Carter, Derrick
 ; APPLICANT: Barth, Brenda
 ; APPLICANT: Douglass, John
 ; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
 ; FILE REFERENCE: 210121.514C1
 ; CURRENT APPLICATION NUMBER: PCT/US02/32727
 ; CURRENT FILING DATE: 2002-10-11
 ; NUMBER OF SEQ ID NOS: 30992
 ; SEQ ID NO 5091
 ; LENGTH: 63
 ; TYPE: PRT
 ; ORGANISM: Propioni acnes
 PCT-US02-32727-5091

Query Match 73.9%; Score 17; DB 1; Length 63;
 Best Local Similarity 25.0%; Pred. No. 4.6e+04;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 |
 Db 56 WAASSTSG 63

RESULT 3

US-09-978-825-5091
 ; Sequence 5091, Application US/09978825
 ; GENERAL INFORMATION:
 ; APPLICANT: Mitcham, Jennifer
 ; APPLICANT: Skeiky, Yasir
 ; APPLICANT: Persing, David
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Maisonneuve, Jean Francois
 ; APPLICANT: Zhang, Yanni
 ; APPLICANT: Wang, Siquing
 ; APPLICANT: Jen, Shyian
 ; APPLICANT: Lodes, Michael
 ; APPLICANT: Benson, Darin
 ; APPLICANT: Jones, Robert
 ; APPLICANT: Carter, Derrick
 ; APPLICANT: Barth, Brenda
 ; APPLICANT: Douglass, John
 ; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
 ; FILE REFERENCE: 210121.514C1
 ; CURRENT APPLICATION NUMBER: US/09/978,825
 ; CURRENT FILING DATE: 2003-01-29
 ; NUMBER OF SEQ ID NOS: 30992
 ; SEQ ID NO 5091
 ; LENGTH: 63
 ; TYPE: PRT

; ORGANISM: Propioni acnes
 US-09-978-825-5091

Query Match 73.9%; Score 17; DB 25; Length 63;
 Best Local Similarity 25.0%; Pred. No. 4.6e+04;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 |
 Db 56 WAASSTSG 63

RESULT 4

US-10-057-498-5091
 ; Sequence 5091, Application US/10057498
 ; GENERAL INFORMATION:
 ; APPLICANT: Mitcham, Jennifer
 ; APPLICANT: Skeiky, Yasir
 ; APPLICANT: Persing, David
 ; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
 ; FILE REFERENCE: 210121.514
 ; CURRENT APPLICATION NUMBER: US/10/057,498
 ; CURRENT FILING DATE: 2001-04-20
 ; NUMBER OF SEQ ID NOS: 29212
 ; SEQ ID NO 5091
 ; LENGTH: 63
 ; TYPE: PRT
 ; ORGANISM: Propioni acnes
 US-10-057-498-5091

Query Match 73.9%; Score 17; DB 26; Length 63;
 Best Local Similarity 25.0%; Pred. No. 4.6e+04;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 |
 Db 56 WAASSTSG 63

RESULT 5

US-09-620-111B-3466
 ; Sequence 3466, Application US/09620111B
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid
 ; FILE REFERENCE: 2750-1070P
 ; CURRENT APPLICATION NUMBER: US/09/620,111B
 ; CURRENT FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 9298
 ; SEQ ID NO 3466
 ; LENGTH: 82
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..82
 ; OTHER INFORMATION: Xaa is any amino acid
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..82
 ; OTHER INFORMATION: Ceres Seq. ID 1327912
 US-09-620-111B-3466

Query Match 73.9%; Score 17; DB 20; Length 82;
 Best Local Similarity 25.0%; Pred. No. 5.5e+04;
 Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 WXXXXXXG 8
 |
 Db 64 WTAATSG 71

RESULT 6

```

US-09-834-366-14776
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (72)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (73)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (93)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; PCT-US01-01321-1364
Query Match 73.9%; Score 17; DB 23; Length 84;
Best Local Similarity 25.0%; Pred. No. 5.6e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 WXXXXXXG 8
| |
DB 13 WSTTSSG 20
RESULT 7
US-60-197-873-14776
; Sequence 14776, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US2.REG
; CURRENT APPLICATION NUMBER: US/09/834,366
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/197,873
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 14776
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-366-14776
Query Match 73.9%; Score 17; DB 23; Length 84;
Best Local Similarity 25.0%; Pred. No. 5.6e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 WXXXXXXG 8
| |
DB 13 WSTTSSG 20
RESULT 7
US-60-197-873-14776
; Sequence 14776, Application US/60197873
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: 81.US1.PRO
; CURRENT APPLICATION NUMBER: US/60/197,873
; CURRENT FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 52153
; SOFTWARE: Patent.pm
; SEQ ID NO 14776
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-197-873-14776
Query Match 73.9%; Score 17; DB 31; Length 84;
Best Local Similarity 25.0%; Pred. No. 5.6e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 WXXXXXXG 8
| |
DB 13 WSTTSSG 20
RESULT 8
PCT-US01-01321-1364
; Sequence 1364, Application PC/TUS0101321
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PCT/US01
; CURRENT APPLICATION NUMBER: PCT/US01/01321
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 2181
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1364
US-09-834-366-14776
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (72)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (73)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (93)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; PCT-US01-01321-1364
Query Match 73.9%; Score 17; DB 17; Length 93;
Best Local Similarity 25.0%; Pred. No. 6.2e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 WXXXXXXG 8
| |
DB 81 WATTSTG 88
RESULT 10
US-10-316-253-138
; Sequence 138, Application US/10316253
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 118
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-316-253-118
Query Match 73.9%; Score 17; DB 29; Length 98;
Best Local Similarity 25.0%; Pred. No. 6.2e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 WXXXXXXG 8
| |
DB 81 WATTSTG 88
RESULT 10
US-10-316-253-138
; Sequence 138, Application US/10316253
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: Patentin version 3.1

```

; SEQ ID NO 138
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-316-253-138

Query Match 73.9%; Score 17; DB 29; Length 98;
Best Local Similarity 25.0%; Pred. No. 6.2e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 81 WATTSTG 88

RESULT 11

US-10-316-253-118
; Sequence 118, Application US/10316253
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 118
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-316-253-118

Query Match 73.9%; Score 17; DB 29; Length 98;
Best Local Similarity 25.0%; Pred. No. 6.2e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 81 WATTSTG 88

RESULT 12

US-10-316-253-138
; Sequence 138, Application US/10316253
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 138
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-316-253-138

Query Match 73.9%; Score 17; DB 29; Length 98;
Best Local Similarity 25.0%; Pred. No. 6.2e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 81 WATTSTG 88

RESULT 13

US-10-316-253-118
; Sequence 118, Application US/10316253
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 118
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-316-253-118

Query Match 73.9%; Score 17; DB 29; Length 98;
Best Local Similarity 25.0%; Pred. No. 6.2e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 81 WATTSTG 88

RESULT 14

US-10-316-253-138
; Sequence 138, Application US/10316253
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 138
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Sus scrofa
US-10-316-253-138

Query Match 73.9%; Score 17; DB 29; Length 98;
Best Local Similarity 25.0%; Pred. No. 6.2e+04;
Matches 2; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 WXXXXXXG 8
|
Db 81 WATTSTG 88

RESULT 15

US-10-316-253-118
; Sequence 118, Application US/10316253